Occupational Allergy to Low Molecular Weight Organic Chemicals: The Role of Structure in Determining Chemical Hazard

Jarvis, J

PhD Thesis, The University of Edinburgh, 1999

Declaration

I, the under-signed, hereby declare that this thesis has been composed by me and the work is my own except where explicitly stated otherwise.

Signed

Date 6/9/1999

Abstract

A number of low molecular weight (LMW) organic chemicals are known to cause occupational respiratory or skin sensitisation. A set of 200 LMW organic skin and approximately 75 respiratory sensitisers were identified by critical appraisal of published case literature. The respiratory sensitisers (asthmagens) were systematically compared in turn with suitable control chemicals and the skin sensitisers using a case-control methodology. The control chemicals were selected from known hazardous LMW organic chemicals for which no reports of respiratory sensitisation could be found.

Several potential methods of differentiating between asthmagens and controls or asthmagens and skin sensitisers (by chemical structure alone) were investigated. These comprised hazardous fragment identification by calculating odds ratios for hazard (HOR's), cluster analysis and the logistic regression analysis. Of these methods the most effective approach was the logistic regression analysis. Using these methods several known or suspected hazardous substructures were confirmed to present statistically significant occupational asthma (OA) hazard. These included isocyanates, acid anhydrides, acrylates and (oligo)-amines. Furthermore, certain sub-structural fragments such as chlorine atoms appeared to provide a protective effect from OA hazard. For differentiating between skin sensitisers and asthmagens it was noted that fragments with carbon double bonded to nitrogen or oxygen atoms were significantly more prevalent in the respiratory sensitisers set.

A predictive model of chemical asthma hazard was created using logistic regression and the model tested on a validation set of chemicals yielded a predictive kappa value greater than **0.7**. This model is available for predictive testing of compounds for asthma hazard via the World Wide Web.

This work demonstrates that simple structural information may, in conjunction with a well designed methodology, be used to identify occupational sensitisers with reasonable reliability.

Abbreviations

A full list of abbreviations is available in appendix D.3

Acknowledgements

I would like to acknowledge the support of The Colt Foundation. Without the Colt Foundation's support this thesis would not have been possible. I also thank my supervisors Raymond Agius and Lindsay Sawyer for their patient and helpful advice. Special thanks to Rob Elton for statistical assistance. Finally, I thank my wife Sinead and my parents for the many years of support!

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Overview

This thesis investigates the degree to which chemical structure can be used to identify occupational sensitising hazard. A null hypothesis which states:

> The potential of a chemical to cause occupational allergic disease cannot be predicted from structure alone.

will be applied. The acceptance or rejection of this hypothesis will be dependent on the validity (or otherwise) of the data, the methods used and hence the results of this study.

Moreover it was an expressed aim that, if at all possible, a working predictive model of occupational chemical asthma hazard should be created. In order to be a predictive system of practical application a model should be simple to use. Throughout, this thesis attempts to elucidate the maximum predictive information from basic topological (connectivity) information of chemical structure. This is based upon

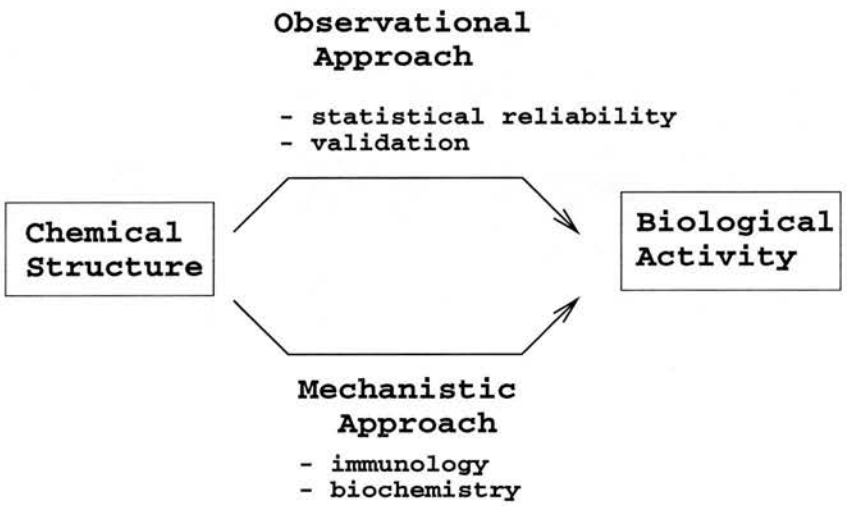


Figure 1.1: Two approaches to structure-activity relationship studies.

the *a priori* assumption that structurally similar molecules will exhibit comparable properties [2].

Two possible approaches to studying the structure-activity relationships (SAR's) are the purely observational approach and the purely mechanistic approach (see Figure 1.1). The mechanistic approach involves forming mechanistic hypotheses based on current knowledge, in this case of the immunology and biochemistry of allergy. The purely observational approach attempts to identify associations between structural data and activity without presuming a particular mechanism or mechanisms. A mechanistic model assumes a causative relationship whereas an observational model need only assume an associative relationship. A well designed observational approach has the advantage of being more objective as it makes fewer *a priori* assumptions.

In this study an observational approach was used and thus the model may be understood in the absence of immunological mechanistic knowledge, although the findings may contribute to mechanistic hypotheses. The use of a purely observational model is open to the criticism that the observed associations may arrive by chance. For this reason the use of statistics to quantify the reliability of associations arising and

the use of a validation procedure to test the final model are crucial if such a non-mechanistic model is to be accepted.

1.2 Epidemiology and the Study of Diseases

The recognition of causal or determinant factors in disease is central to the way illness is viewed and treated. People would associate their sickness with something they had done wrong. In Biblical times disease was seen as a consequence of sin. Indeed when healing the paralytic Jesus said...

> ...your sins are forgiven. (Mark,2:5.)

In the the 19th Century Louis Pasteur's ideas of disease being due to infectious agents prevailed. So when Casimir Funk suggested that the disease beri-beri was due to a "vitamine" deficiency his ideas were not immediately accepted [3]. However it was becoming apparent that astute observation of case histories and lifestyle of the sick could give clues to the determinants of disease. Perhaps the finest 19th Century example of a methodical approach to identifying disease determinants was John Snow's work *On the Mode of Communication of Cholera* [4].

Epidemiology is the study of diseases in populations [1, p1]. It is however more than that; it is the systematic study of determinants of diseases and the distributions of diseases within populations. In the forward to Richardson's book reprinting John Snow's work [4], Frost gives the following description of the discipline:

> Epidemiology at any given time is something more than the total of its established facts. It includes their orderly arrangement into chains of inference which extend more or less beyond the bounds of direct observation.

The implication is that through an ordered approach the discipline of epidemiology makes apparent those associations that would be hidden from cursory view.

Epidemiology is founded on two fundamental assumptions about the nature of disease: firstly, it assumes that disease does not occur randomly; and secondly, it assumes causal (and therefore preventive) factors are responsible for disease. Often these factors are in the environment.

Criteria for determining disease causation are important. In theory a causal factor should be both *necessary and sufficient* to cause the effect. In reality what is deemed 'causal' will depend on current knowledge - at the time of Snow a particular water supply was deemed causal without knowledge of underlying micro-organism, *Vibrio cholerae*. Associations alone may arise by chance so it is important to have criteria for determining how the associative relationship might reflect a causal relationship. Table 1.1 lists some of the criteria which should be applied to determine causation taken from Hennekens and Buring [1]. These criteria are not absolute, if they are not all fulfilled it is possible a causal relationship still exists. An outcome that is all or nothing[1] (i.e. occurs at a threshold) and irreversible will fail to satisfy the temporal and the dose-response criteria.

Case-control studies are frequently used in epidemiological research because of their relative simplicity and low cost. They often represent the only practical methodology for studying rare diseases. The present study can be said to implement an approach analogous to a case-control study.

Usually a case-control study involves the identification of subjects with the disease and control subjects selected to be representative of the undiseased individuals. The groups are then compared in order to identify differences in perhaps environment or lifestyle exposures which

[1]Death for example!

Criteria
Is there a valid statistical association?
- is the association likely to be due to chance?
- is the association likely to be due to bias?
- is the association likely to be due to confounding?
Can this valid statistical association be judged as cause or effect?
- is there a strong association?
- is there biologic credibility to the hypothesis?
- is there consistency with other studies?
- is the time sequence compatible?
- is there evidence of a dose-response relationship?

Table 1.1: Framework for the interpretation of an epidemiologic study. (After Hennekens and Buring [1, Page 45, Table 3-1], though this is based on Bradford Hill's original observations on Statistical Evidence and Inference [5])

may be associated with the disease occurring. The method is prone to bias in selection (of cases and controls) from the population as a whole and bias in the identification of exposures for each of the two groups. The case-control methodology is not suitable for identifying temporal factors in disease aetiology but has the advantage that is considerably cheaper than long term monitoring of a population that is required for a cohort study.

Occupational Health And Disease

The types of factors proposed as responsible for disease have varied across time and culture. Postulated causes for disease may include divine retribution for one's sins or the effects of evil spirits. Such explanations were often all that was left if no obvious external cause was available. It is perhaps inevitable (given the proportion of adult life spent pursuing an occupation) that some disease will occur as a consequence of a person's employment. Indeed it is not difficult to imagine the variety of occupational hazards that accompanied the development

of mankind through the various stages of history. Socrates is credited with the observation

> 'What are called the mechanical arts...
>
> ...damage the bodies of those who work at them...'[6]

but to the ancient Greek 'citizens' occupational disease was of little importance because it did not directly affect them [6]. Mining, one of the most hazardous of occupations is also amongst one of the oldest. In 1556 Georgius Agricola's work *De Re Metallica* described in one of its volumes the diseases, accidents and hazards particular to mining [6].

The landmark text in occupational respiratory disease was written by Bernardino Ramazzini *(De morbis artificum diatriba)* in 1713 [7, 8]. According to Butcher and Salvaggio [8], Ramazzini's was the first published description of dyspnoea after inhalation of organic dust. Ramazzini is also described as the Father of Occupational Medicine [6] since it was he who suggested an addition to the questions Hippocrates recommended physicians put to their patients. That addition was to enquire after the patient's occupation.

1.3 Allergic Disease

Immunology of Allergy

Allergic reactions are caused by an inappropriate response of the subjects immune system. Allergy is defined in the dictionary [9] as "a hyper-sensitivity to a substance that causes the body to react to any contact with that substance". A more immunological definition is one that states hyper-sensitivity occurs when "immune reactions are out of all proportion to the damage that may be caused by a pathogen"

[10]. Similarly sensitisation is defined as the process "to make or become sensitive; to render (an individual) sensitive to a drug, allergen, etc.." A subject who is allergic/sensitised to a particular agent - the allergen/sensitiser - will exhibit an adverse reaction to that agent which is greatly in excess of the reaction that would be observed when a non-allergic subject undergoes the same exposure.

Diagnosis of allergy requires a demonstration of hyper-sensitivity. To be a true allergy the allergen must evoke in the subject a response which is not seen in the normal population exposed to the same degree. Furthermore this reaction must be demonstrated to be specific and not the result of non-specific hyper-reactivity - the condition in which the subject reacts to agents to which they may never previously have been exposed. (In practice the distinction between non-specific and allergic hyper-reactivity may not be clear cut [11, 12].)

Occupational allergy covers a spectrum of diseases including dermatitis, conjunctivitis, rhinitis, asthma and in the most severe cases anaphylaxis [13]. The pathway by which allergy occurs in an occupational setting is typically contact or inhalation. Allergy may be caused by sensitivity to a vast range of chemical agents. According to Hodgson *et al.*, the prevalences of occupational dermatitis and asthma in the U.K. (based on self-reported symptoms) are 54,000 and 20,000 respectively [14]. Occupational allergy therefore accounts for significant workplace morbidity. Consequently a means of identifying novel hazardous chemicals represents an opportunity to prevent further morbidity.

Immune Defences

The body can be considered to be protected from attack by infectious agents by three forms of defence. The most obvious are the physical barriers, the skin, the epithelial linings of the lung and the gut. These barriers are highly effective. The second form of defence is the innate

immune response. This is the response to cell surface markers common to many pathogens. These responses are mediated by the complement system. The third form of defence is the adaptive immune response - a mechanism by which the body learns to identify and destroy pathogens. On re-exposure to the same pathogen the response is faster and more effective. It is this third form of defence responding inappropriately that results in allergic disease.

Types of Adaptive Immune Responses

Adaptive immune responses are classified into one of four response types [15, 10].

- Type I Reactions - Anaphylaxis - *asthma*
- Type II Reactions - Antibody-mediated cytotoxicity
- Type III Reactions - Immune complex disease
- Type IV Reactions - Cell-mediated immune reactions - *dermatitis*

Of these, Types I and IV are the most important with respect to allergy and will be described in more detail in the following sections. Type II reactions include auto-immune responses. Blood cells are an example of the cell targets of Type II reactions which result in the lysis and phagocytosis of these cells. Type III reactions occur when antibody-antigen complexes form at a critical intermediate size (smaller, soluble complexes can readily be cleared and larger complexes can be removed by the reticulo-endothelial system). The intermediate complexes however aggregate around the basement membranes triggering inflammatory mediator release. An example of a Type III reaction would be the serum sickness observed when a heterologous serum transfusion is

given. Types II & III (hyper-sensitivity) reactions are not believed to be relevant in allergic reactions.

Although the immune system has some innate (non-specific) defence mechanisms (such as complement formation and phagocytosis) those mechanisms involved in allergy are the result of adaptive immune responses. In particular, this includes the recruitment of antibodies (also known as immuno-globulins (Ig's)) which can specifically recognize foreign molecules (antigens) and bind to them. Antibodies are proteins which may be schematically represented as 'Y'-shaped molecules. The arms of the 'Y' are the binding regions, two per molecule, and the stem is known as the *constant* region. Once antibody has bound antigen, that antigen is marked as an appropriate target for an immune response. This response may include the stimulation of memory cells, the release of inflammatory mediators or the recruitment of phagocytic cells.

Many of the agents that cause occupational asthma are low molecular weight chemicals. Low molecular weight chemicals may not be true antigens. More typically they may be haptens - molecules which in isolation are too small to be immunologically recognized. Haptens only become recognized by antibodies after conjugation to another molecule, typically a host protein (see Figure 1.2). The hapten-protein conjugate is recognised as foreign by the host immune system and an immune response results. Various studies [16] have shown that haptens need to bind in relatively large numbers to host proteins to produce effective antigenic conjugates. It has been suggested that OA is mediated through a Type I immune response [17] and while this may be true for the majority of cases there may be exceptions [18].

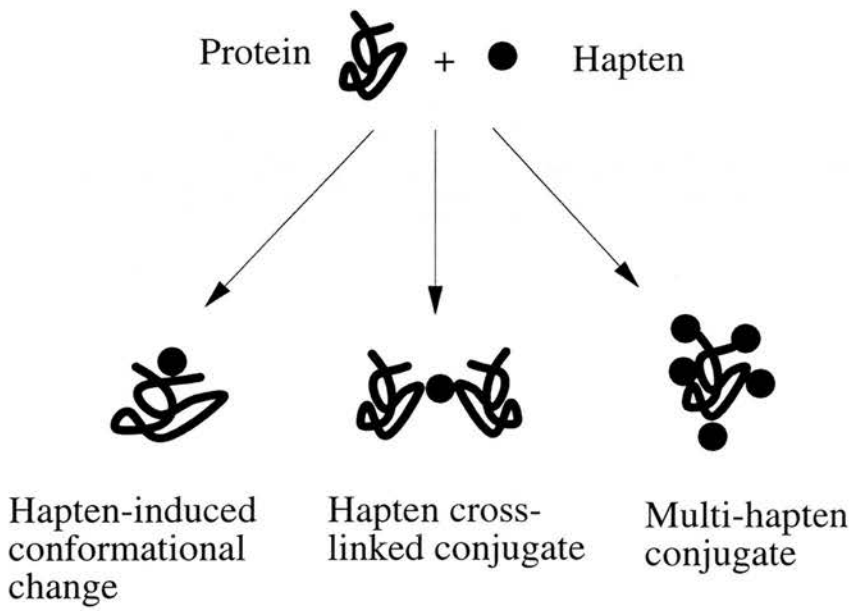


Figure 1.2: Antigen formation.

Type I (Hyper-sensitivity) Reactions - Anaphylaxis

Usually the term allergy is used to refer to Type I (immediate hypersensitivity) immune reactions. Type I reactions are extremely rapid, an example being the immediate wheal and flare seen on the skin of hay fever suffers when scratch tested with pollen extracts. Type I reactions account for the majority of cases of asthma. The term anaphylaxis is also used to describe Type I reactions, particularly the more severe ones in which there is constriction of smooth muscle, vasodilation and a very real risk of death.

Type I reactions are mediated by the antibody/immunoglobulin (Ig) IgE. Antibodies (immunoglobulins) are protein molecules which possess an ability to recognise and bind to other molecules. Reaginic antibodies (reaginins) are cytophyllic antibodies produced in response to allergen. The term reaginic is more commonly encountered in the older literature. Reaginic antibodies refer to the E class of immunoglobulin, IgE and to a lesser extent the G class, IgG.

IgE antibody is found in serum and attached to $FC_{\varepsilon}RI$ receptors on mast cells. These receptors bind to what is known as the constant region of the IgE antibody (the stem of the 'Y'). In serum the half life of IgE is only a few days but IgE bound to $FC_{\varepsilon}RI$ receptors on mast cells may last for many months [10]. Normal levels of serum IgE are in the range 10 to 200 IU/ml (median $\simeq$100 IU/ml) [10]. In atopic individuals levels of IgE are elevated, in particular levels of IgE specific to the allergen(s) to which the subject is allergic.

During the initial exposure(s) antigen is processed and presented by antigen-presenting cells (APC's). APC's present the processed antigen to $CD4^{+}$T-Helper_2 (T_{H_2}) cells which release cytokines, principally interleukins 4 and 13 (IL4 and IL13). ($CD4^{+}$ indicates that the cell expresses the CD4 cell-surface marker.) These cytokines stimulate B cell proliferation resulting in the production of large amounts of IgE specific to the antigen. The resulting IgE binds to mast cell IgE receptor $FC_{\varepsilon}RI$. On subsequent exposure the antigen combines with these IgE on the $FC_{\varepsilon}RI$ receptors on mast cells stimulating degranulation of the cell. On degranulation the mast cell releases histamine and other cytokines which trigger the inflammatory response.

Following antibody binding to antigen, the resulting combination is then capable of triggering membrane receptors on mast cells. This results in the release of mediators including histamine, leukotrienes, prostaglandins, bradykinins and eosinophil chemotactic factors. The antibodies involved are typically of the IgE and IgG classes.

Type IV (Hyper-sensitivity) Reactions - Cell-mediated immune reaction

Type IV Reactions are known as cell-mediated responses or delayed hyper-sensitivity. While Type I reactions occur very rapidly following

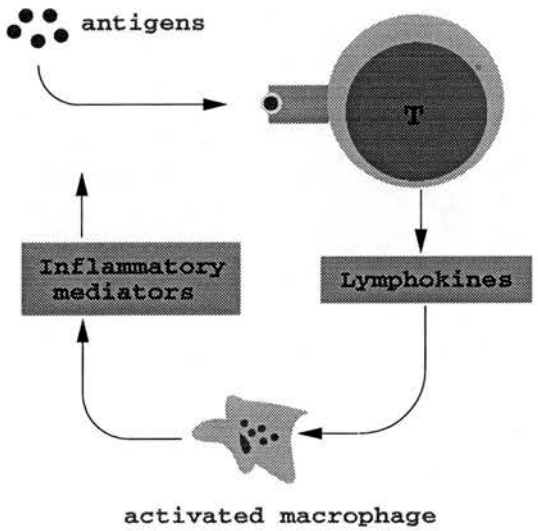


Figure 1.3: Mechanism of the Type IV (delayed hyper-sensitivity) reaction.

antigen exposure Type IV reactions can take up to 72 hours to appear [10]. Delayed hyper-sensitivity reactions include the skin reactions resulting from intra-dermal injection of allergen. The interaction between the thymus-derived lymphocytes (T-cells) and the allergen to which they are specific results in the release of lymphokines, some of which are cytotoxic to specific target cells (see Figure 1.3).

The sensitisation process in humans can take 10-14 days [10]. Langerhans cells in the epidermis internalise hapten-protein conjugates and migrate to the local lymph node the via afferent lymphatics. Here they present antigen in conjunction with interferon-γ (IFN-γ) and tumour necrosis factor-β ($TNF-\beta$) [19]. The stimulation of IFN-γ is typical of a T_{Helper_1} (T_{H_1}) cell response and is characteristic of skin sensitisation [20, 17]. Indeed IFN-γ antagonizes the IgE production associated with T_{H_2} cell activity [17].

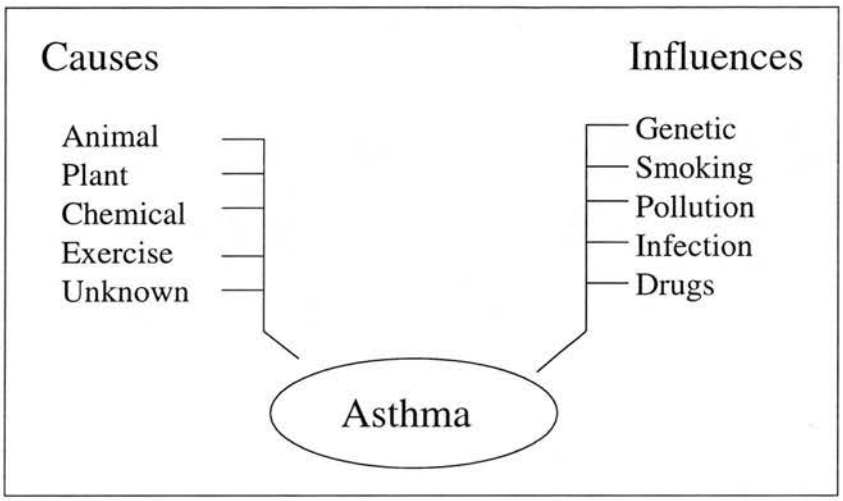


Figure 1.4: Asthma can be influenced by a number of factors.

Occupational Asthma

Occupational Asthma (OA) is asthma caused by sensitisation of the respiratory tract to workplace chemicals. OA is one of the two forms of workplace asthma, the other being work-aggravated asthma, in which pre-existing asthma is made worse by exposure to triggers in the workplace. Indeed any case of asthma could be influenced in a number of ways (see Figure 1.4) and consequently there may be difficulty in diagnosing OA as distinct from asthma caused outwith the workplace.

OA is more readily defined by its symptoms and underlying pathophysiology than by its immunology. A wide variety of symptoms such as cough, chest tightness, wheeze and nasal symptoms may accompany sensitisation. The factors underlying the main symptoms are reversible broncho-constriction and excess mucus production. There is marked variation in symptom presentation. The temporal pattern of the disease also varies, some cases showing an immediate onset of symptoms following exposure, whereas in others the symptoms occur several hours later[1] . This variation is seen both between different

[1]When symptoms do occur several hours after the exposure this can make diagnosing occupational asthma and identifying the responsible agent extremely difficult.

chemical causes and between different patients sensitised to the same chemical. The term OA has therefore been said to cover a spectrum of respiratory disorders [21].

The diagnosis of asthma is performed on the basis of observed symptoms and measurements of lung function. The clinical diagnosis of asthma is often arrived at by these methods without any identification of the causative agent. Asthma is generally diagnosed as intrinsic or extrinsic. One of the simplest diagnostic measures is peak expiratory flow rate (PEFR) monitoring. This is simply the maximum rate a person can expel air from his lungs. It is measured by a peak flow meter. A more informative diagnostic measure is the forced expiratory volume in one second (FEV_1) which is measure using a spirometer. It is often desirable to identify whether a reaction is antibody-mediated. Various methods exist for measuring both non-specific (total) and specific serum antibody levels. Assays of specific antibody concentrations (using for example ELISA techniques [22]) are frequently used because they are readily available through diagnostic laboratories.

There is a considerable variety of chemicals that cause OA (see p30) so there may also be problems identifying the causal agent in an environment containing mixtures of chemicals. Indeed the gold standard[1] diagnosis of OA to particular chemical can only be achieved by blind challenge inhalation with pure compound whilst having controls to eliminate airway hyper-sensitivity as a cause. Such tests are not trivial, require specialist equipment and should always be carried out within a hospital. Some studies also include a challenge of a non-sensitised control asthmatic exhibiting similar pre-test airways reactivity to the patient studied [23].

Asthmatic responses may be immediate, delayed or dual. The immediate response, which may be seen in isolation as in Figure 1.5, is believed to be due to a narrowing of the airways by the contraction of the

[1]The term given to describe the best available diagnostic test.

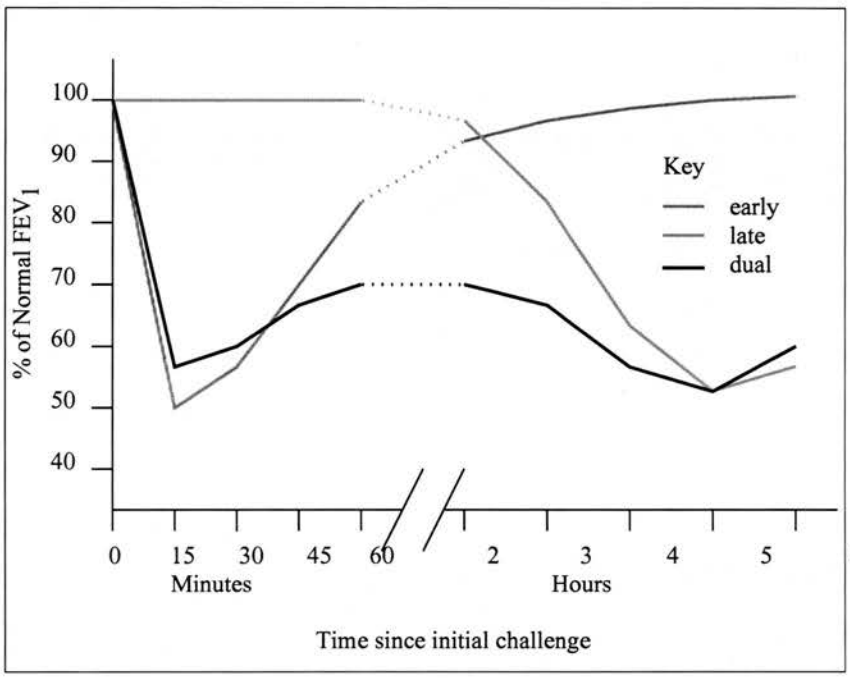


Figure 1.5: Change in FEV_1 For Early, Late and Dual Asthmatic Responses.

Schematic Representation of change in FEV_1 for early, late and dual asthmatic responses. (After Cartier and Malo, 1993 [25].)

smooth muscle walls in response mainly to released histamine from mast cells. It is usually indicative of a Type I immune reaction (see p24). The response occurs quickly, in some cases in seconds, in others up to 15 minutes after challenge. Since the early response is primarily due to smooth muscle contraction it responds well to recovery treatment using β-adrenergic drugs such as salbutamol [24].

A dual response type (see Figure 1.5) is commonly seen with classical Type I (IgE-dependent, atopy-related, see p24) immediate hypersensitivity asthma response. An initial fall in lung function (as measurable by FEV_1) is observed within minutes of challenge. This is followed by an improvement of FEV_1 though not necessarily a return to pre-challenge value. Several hours later a second fall in FEV_1 is observed. The cause of the first fall in FEV_1 is usually the constriction of the circular smooth muscle of the bronchi walls. The late drop in FEV_1

is typically the result of an inflammatory response resulting in mucus plugging of the airways.

A delayed response (see Figure 1.5) typically occurs several hours after exposure to the causative compound. This must often preclude early diagnosis of OA because symptoms occur after the subject returns home from work thus the correct association between the symptoms and the chemical exposure is not made. The delayed response type alone is rarely described in non-OA but is seen in cases of asthma due to aldehydes, amines, acrylates, anhydrides and isocyanates. A delayed type response may be indicative of a Type IV immune response (see p25) however it may also indicate a transport latency between chemical challenge and the chemical reaching its site of action. Interestingly, Durham *et al.* [26] notes that

> ... increases in airway responsiveness precede the late asthmatic response, occur independently of changes in airway caliber, and correlate with the magnitude of the subsequent late response.

Consequently the delayed response may represent the latency of transport and triggering of an inflammatory response.

Although Figure 1.5 indicates typical temporal profiles of FEV_1 each individual case is different. Occasionally unusual response profiles may be noted such as early-late, progressive, square-waved or prolonged-immediate [25].

Classes of Low Molecular Weight Occupational Asthmagens

A comprehensive classification of occupational asthmagens may include mammal-derived antigens, insects, flour, tea, wood dusts, vegetable

gums, laboratory and commercial enzymes, diisocyanates, anhydrides, aliphatic amines, ethanolamines, heterocyclic amines, aromatic amines, fluxes, metals, drugs, synthetic materials and a selection of other chemicals. This diversity of chemical asthmagens produces a great heterogeneity of symptoms, such that OA has been described as a spectrum of disorders [21]. The LMW organic asthmagens can be classed as follows:

Isocyanates

Isocyanates are esters of isocyanic acid (H-N=C=O). Polyfunctional isocyanates are used in the manufacture of polymeric foams, fibres, coating and solid elastomers because they react readily with many functional groups with a limited amount of undesirable by-product formation. The normal reaction of isocyanates involves the addition of an 'active' hydrogen, that is, one replaceable by sodium (Figure 1.8).

Isocyanates find a wide variety of applications from use in spray paints to the manufacture of mechanical goods, industrial tyres, coated fabrics, shoe products, wire and cable. They are extremely reactive chemicals which have several other toxic effects in addition to being sensitisers. Isocyanates are the most commonly cited and probably the most potent cause of OA, accounting for in excess of 20% of all cases of OA reported in the U.K. [27]. One recorded OA fatality was due to exposure to the exposure of a sensitised worker to paint containing just 0.15% toluene di-isocyanate (TDI, Figure 1.6) [28]. It is worth noting that TDI generally comes as a mix of the 2,4- and 2,6- isomers in a ratio of about 4:1 [29].

Several other di-isocyanates have been described as causing OA including hexamethylene di-isocyanate (HDI, Figure 1.7) [30, 31], isophorone di-isocyanate [32], diphenylmethane di-isocyanate (MDI) [33], 1,5-naphthalene di-isocyanate (NDI) [34]. Respiratory symptoms have been observed in

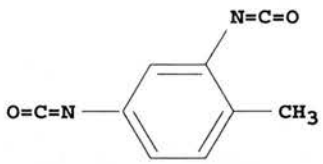


Figure 1.6: Toluene-2,4-di-isocyanate.

Figure 1.7: Hexamethylene di-isocyanate.

Gulf War veterans due to exposure to TDI present in the chemical agent resistant coatings used to paint military vehicles [35].

It is notable that all the reported causes of isocyanate asthma are due to di- or poly-isocyanates. This may reflect the more widespread use of and therefore exposure to these compounds compared with mono-isocyanates (for reasons detailed below) but may also be a reflection of a fundamental requirement for bi- or poly-functionality in the immunochemistry of isocyanate-induced asthma. However until a detailed study of workers exposed to mono-isocyanates is undertaken this will remain just a hypothesis.

It is possible that some individuals are more prone to developing asthma to isocyanates. Savolainen [36] states that α_1-anti-trypsin carriers may be at a greater risk of developing di-isocyanate asthma. The study was based on only a small number of subjects so this result remains to be confirmed.

Asthma due di-isocyanates can persist despite negligible workplace exposure [37] and continued exposure leads to worsening symptoms [38]. The symptoms that accompany di-isocyanate respiratory sensitisation may include dyspnoea, wheezing, chronic cough, phlegm and rhinitis [39].

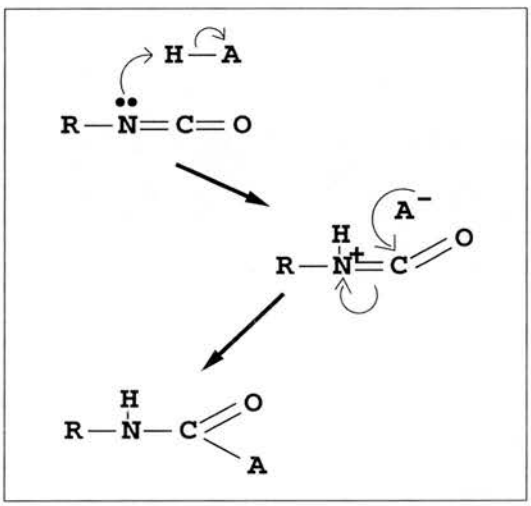


Figure 1.8: Isocyanates react by an addition mechanism.

Acid Anhydrides

Acid anhydrides are formed when two carboxyl groups react in such a way that a water molecule is released. They are named after their parent acid. Acid anhydrides are used as curing agents in epoxy resins [40] Venables describes in detail the uses and toxicity of acid anhydrides [40]. Amongst the compounds she cites as capable of causing OA are phthalic anhydride [41, 40, 42] (see Figure 1.9), tetrachlorophthalic anhydride [43, 44, 45, 40, 46], trimellitic anhydride [42, 47, 40, 48, 40, 49], maleic anhydride [50], and pyromellitic di-anhydride [40]. Methylhexahydrophthalic anhydride [51], hexahydrophthalic anhydride [48, 47, 40, 52], himic anhydride [49] and methyltetrahydrophthalic anhydride [53, 54, 55] have also been reported to cause respiratory sensitisation.

Acid anhydrides are amongst the most common chemical causes of OA [27]. They have a variety of adverse effects [40, 52]. At least one death is reported which may have been in part due to acid anhydride-related asthma [45].

Although acid anhydrides can and do have irritant effects they are believed to cause asthma by acting as haptens, reacting with endogenous

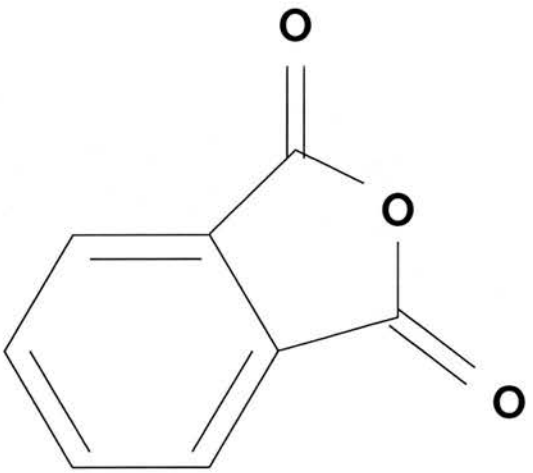


Figure 1.9: Phthalic anhydride

protein and eliciting a Type I (IgE-mediated) hyper-sensitivity reaction [43, 56, 44]. There is evidence that the IgE antibodies characteristic of sensitisation may persist for several years after exposure [45]. Enzyme-linked immuno assays can be effective in identifying workers sensitised to anhydrides [42]. A hapten density in a range 10-25 mol/mol serum albumen appears to be optimal for allergic reactivity [57] of acid anhydrides. Other affects noted from anhydrides include haemorrhagic rhinitis, reported due to to exposure to hexahydrophthalic anhydride[52].

The Tree Acid Resins, Abietic and Plicatic Acid

Abietic and plicatic acids are found in colophony (rosin) from pine trees and from cedar respectively. They are known asthmagens. There may be other asthmagens present in wood extracts but none are as well characterized as these two chemicals.

Colophony is a natural substance obtained from pine trees [58]. It has a complex chemical composition of which approximately 90% consists of di-terpene resin acids. Colophony is used in soldering fluxes to remove surface corrosion (oxides, sulphides or chlorides) [58]. The hot

solder then destroys the flux thus allowing a good contact with the underlying metal. Most cases of OA due to colophony arise from its use as a soldering flux but Burge *em et al.* [59] report a case in a man who had been working with a bitumen mixture containing colophony. In a respiratory challenge test he suffered a dual reaction to colophony at room temperature.

Plicatic acid is believed to be the constituent of both eastern white cedar (Thuja occidentalis)[60] and western red cedar (Thuja plicata)[61, 62] that is responsible for the respiratory sensitisation of workers exposed in the timber and sawmill industry.

Reactive Dyes

Reactive dyes are another potent class of sensitiser. Imperial Chemical Industries first marketed a reactive dye for cotton in 1956 and since then many more reactive dyes have become available. The popularity of reactive dyes stems from the availability of bright colours and the permanence with which they dye textiles. Reactive dyes covalently attach to hydroxyl and amino groups on the cloth. This is facilitated by reactive groups such as heterocyclic aromatic halogens or reactive groups on aliphatic side chains. A further important feature required for efficient dying is that there must be hydrophilic groups if water is to be used as the dying solvent. The dye colour is due to a light absorbing system, usually an azo adjacent to aromatic rings.

OA due to reactive dyes typically involves a Type I immune reaction and invariably, when sort, increased levels of IgE can be demonstrated. In several cases of OA to reactive dyes, IgE has been demonstrated [63, 64, 65, 66]. Romano *et al.* describe a case of immediate (therefore probably type 1 immune response) in a worker exposed to the bromo-acrylic dye, Lanasol Yellow 4G [67]. Noferi describes five cases of asthma due to the soluble azo dye Direct Black [68]. These reports in

conjunction with animal studies strongly suggest that OA due to reactive dyes is a classic type 1 immunological response. This is supported by the observation that the asthmatic response is immediate following a respiratory challenge [63]. There is also evidence that reactive dye induced OA may occur with an absence of non-specific bronchial hyperreactivity. Reactive dye asthma has been implicated in an OA fatality [69].

Some non-reactive dyes and colouring agents known to cause asthma include methyl blue (a component of ECG ink) [70, 71] and carmine (a natural red dye and food colouring agent) [72, 73].

Aldehydes

Formaldehyde and glutaraldehyde are amongst most commonly used LMW chemicals reported to cause asthma. Formaldehyde is widely used in pathology to preserve and fix specimens. It has powerful disinfectant properties. It is used in industry in the formation of urea-formaldehyde thermosetting resins [74]. Similarly, glutaraldehyde is used as a disinfectant (particularly in disinfecting surgical equipment [75, 76]) and is also used industrially for tanning leather [74].

The case for formaldehyde being a respiratory sensitiser is controversial [77]. A study amongst medical students (exposed during anatomy classes to formalin-preserved cadavers) found little evidence to suggest respiratory sensitisation was occurring [78]. It should be noted however that formaldehyde has been shown to produce asthmatic symptoms when a 'sensitised' worker was exposed to formaldehyde in a challenge test [79]. Burge *et al.* (1985) [80] noted a significant asthmatic reaction in a patient challenged with 1% formaldehyde solution. The same patient experienced rhinitis and watery eyes 10 hours after exposure to 0.1% formaldehyde. Clearly, formaldehyde sensitisation does occur in some individuals. From a mechanistic standpoint

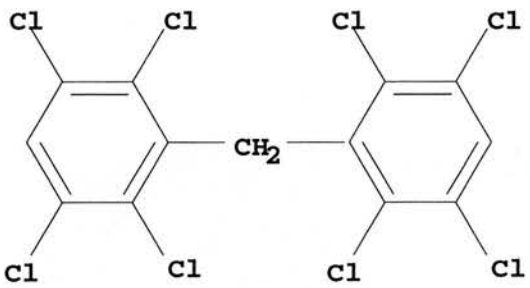


Figure 1.10: Hexachlorophene - a germicidal additive to soaps.

formaldehyde does not appear to elicit the standard IgE type I immunological typical of other respiratory sensitisers [77]. This may indicate that formaldehyde asthma is due to an irritant effect (as suggested by Hilton *et al.*, 1996) [77] or an as yet poorly understood immunological process.

Other Cleaning and Sterilising Agents

Burge and Richardson report a case of a pharmacist sensitized to the cleaning agent lauryl dimethyl benzyl ammonium [81][1]. Lauryl dimethyl benzyl ammonium chloride is one of a class of biocides with general formula 'alkyl' dimethyl benzyl ammonium chloride. A report by Innocenti describes a case of OA due to benzalkonium chloride but does not sufficiently speciate the compound [82]. Chloramine (chloramine-T) is used as a sterilising agent in industries such as brewing. There have been several reports of it causing OA [83, 84, 85]. Chlorhexidine, a known skin sensitizer, caused asthma in a nursing auxiliary exposed to the compound in a chlorhexidine and alcohol spray [86]. The topical antiseptic/disinfectant hexachlorophene [87] which is used as a germicidal additive in soaps has also caused asthma in a nurse (see Figure 1.10). Finally, a detergent additive, sodium iso-nonanoyl oxybenzene sulphonate [88] caused asthma in a laboratory technician.

[1]The same paper also mentions that the pharmacist had his first asthma attack after a spill of chloroxylenol over 20 years earlier[81].

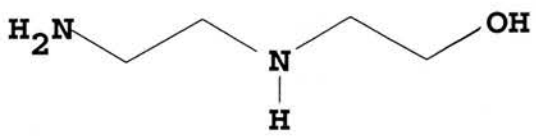


Figure 1.11: Aminoethyl-ethanolamine

Ethanolamines

Ethanolamines are widely used in the production of soaps, detergents, spray paints, cosmetic formulations and shampoos due to their emulsifying properties. They are also used as lubricant additives, plasticizers and corrosion inhibitors [89]. Ethanolamines are also found in soldering fluxes [90].

There have been a number of reported cases of OA due to ethanolamines: ethanolamine [91, 89], aminoethyl-ethanolamine [92, 90] (see Figure 1.11), dimethyl ethanolamine [23], triethanolamine [89] and 2-diethylethanolamine [93]. 3-(Dimethylamino)propylamine, a volatile component of an epoxy resin used in ski manufacture, has also been reported to cause symptoms not incompatible with a diagnosis of OA [94]. Interestingly there is an absence of reported cases of asthma due to ethanolamines in cosmetics manufacture. This may indicate that the differences in method of ethanolamine employment are critical to health outcomes. One such difference may be the temperature at which the ethanolamines are used.

Other Amines Known to Cause Asthma

These include ethylene diamine [95, 96, 91], piperazine / piperazine dihydrochloride [91, 97, 98], hexamethylene tetramine [91], N-methyl morpholine [99], paraphenylene diamine [100]and possibly 4,4'-diaminophenylmethane [101].

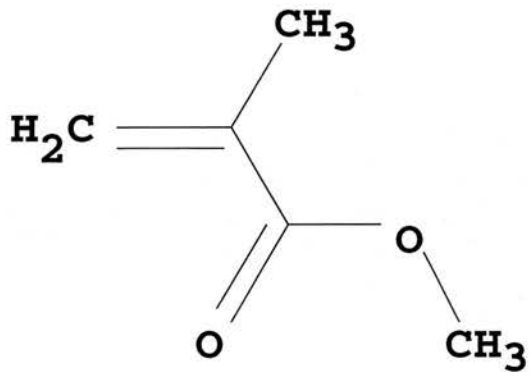


Figure 1.12: Methyl-methacrylate is found in super-glues.

Acrylates Known or Suspected to Cause Asthma

Acrylates and cyanoacrylates known to cause OA include methyl-methacrylate [102, 103] (see Figure 1.12), methyl-cyanoacrylate [104, 103, 105] and ethyl-cyanoacrylate [103]. Nakazawa describes a case of (probably IgE-mediated) OA due to "alkyl"-cyanoacrylate in super-glue but does not specify which alkyl group was present [106]. Chan *et al.* report a case of asthma due to "cyanoacrylate" but again they do not specify the exact chemical [107][1].

Methylmethacrylate is used as bone cement in surgical operations [105], cyanoacrylates are commonly found in super-glues [106]. It has been noted that ambient air concentrations of acrylates are dependent upon humidity. In high humidity the water vapours "mop" up cyanoacrylate vapours thereby neutralising the hazard [103]. Chan *et al.* note a low incidence of cyanoacrylate-induced asthma in Singapore is probably due to the high humidity there [107]. Savonius *et al.* [108] cite a number of suspected cases (several requiring further confirmation). These include asthma apparently due to acrylic acid, methacrylate, isobutyl-methacrylate, trihydroxy-methyl-propyl-triacrylate, hydroxy-propyl-acrylate and methyl-cyanoacrylate.

[1]It is unfortunate that so many case reports, despite performing challenge tests, fail to identify the exact chemical responsible.

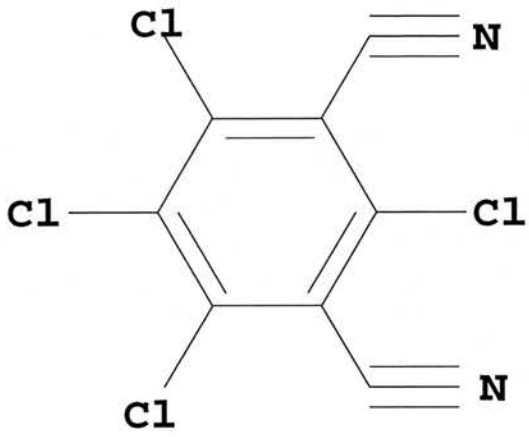


Figure 1.13: Tetrachloroisophthalonitrile - one case of asthma reported due to this fungicide.

Fungicides and Insecticides

Fungicides and insecticides have on rare occasions been reported to cause OA. The fungicides tetrachloro-isophthalonitrile [109] (see Figure 1.13) and captafol (Difolatan) [110]; and the insecticides phosdrin (also known as mevinphos) [111], fenthion [112] and dichlorvos [112].

Drug and Drug Intermediates

A large number drugs and drug intermediates have been linked to asthma both following administration therapeutically and by inhalation by workers handling these chemicals. Amongst the cases of OA the subjects have usually been workers exposed to large quantities of the drug during its manufacture. On occasions allergic reactions have been noted in medical personnel charged with the task of administering these drugs to patients.

The most common pharmaceutical products linked with OA are β-lactam antibiotics. These include benzyl penicillin [113], 6-amino penicillamic acid [113], 7-amino cephalosporanic acid [114], ampicillin [113], cephalexin [13] and amoxicillin [115]. The β-lactam compounds are not the

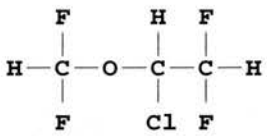


Figure 1.14: Enflurane is an inhalation anaesthetic.

only antibiotics to be implicated, cases of OA due to the macrolide antibiotic spiramycin [116, 117, 118] and tetracycline [119] have also been reported.

In addition several other compounds which are either medicinal agents or intermediate chemicals in the manufacture of medicines have been reported to cause asthma. The include the anti-hypertensive drug methyl DOPA [120], the anti-ulcerative drug cimetidine [121], the opiate morphine [122], MM22383 [123], dioctyl phthalate [124], phenylglycine acid chloride [125], hydralazine (an anti-hypertensive drug) [126], sulfathiazole (used in the manufacture of the bacteriostatic sulphanilimide) [127], isonicotinic acid hydrazide (an anti-tuberculosis drug also known as isoniazid) [128], penicillamine [129] and glycyl compound [130]. There has also been a report of changes of self-recorded air flow measurements in an anaesthetist repeatedly exposed to enflurane[1] [131]. Finally there is a report that salbutamol [132] and another drug used in the treatment of asthma, aminophylline [133] have been reported to cause asthma.

Other Chemicals

Other chemicals reported to cause asthma include styrene (used to make polystyrene plastics) [102, 134, 135, 136], Δ-3-carene (found in rubber gloves) [137], azobisformamide (a blowing agent also known as azodicarbonamide) [138, 139], Pauli's reagent [140], tetrazene (used in detonator manufacture) [141], furfuryl alcohol (in a furan-based binder

[1]The paper descibing asthma due to enflurane was missed and so this compound was not included in the study until it was entered as a validation compound.

system) [142], amprolium hydrochloride (a poultry food additive) [143], ethylene oxide [144, 145] and possibly hydroquinone [146] and methionine [146]. The majority of these (styrene excepted) have only been documented as causing asthma in one or two cases.

Occupational Contact Dermatitis

Occupational contact dermatitis (OCD) is the skin sensitisation as a result of dermal contact with workplace chemical agents. Many hundreds of chemicals have been reported to cause cause OCD, though a proportion of these will cause irritant rather than allergic dermatitis. The symptoms of dermatitis may be localized to the point of contact or in severe cases may spread to include other areas of skin. The area may show erythema (reddening), be pruritic (itchy), have maculo-papules (small raised, circumscribed lumps of infiltrating inflammatory cells), be inflamed and in severe cases the stratum corneum may be cracked.

Irritant dermatitis is due to acute effects of the chemical on the skin. Irritant dermatitis does not imply any sensitisation and is not of an allergic nature. Allergic contact dermatitis implies that sensitisation does occur and that re-exposure will evoke a more pronounced response. It is possible for a chemical to be a cause of both irritant and an allergic contact dermatitis. It is also possible that chemicals with the potential to sensitize only do so in the presence of a secondary irritant substance. Such a secondary adjuvant is used in immunological studies of allergy to promote and accelerate the sensitisation process. Mechanical compromise of the skin or constant 'wet work' can also facilitate the onset of contact dermatitis to chemical agents. Dermatitis can also be aggravated by the presence of pathogens such as *Staphylococcus aureus* [147].

The two phases to the presentation of allergic contact sensitisation are induction and elicitation. Induction is the phase during which sensi-

tisation occurs. Elicitation is the subsequent eczematous reaction to re-exposure to the substance.

A wide variety of occupations present an occupational contact dermatitis hazard. In addition to the manufacturing industries in which chemicals are made and used personnel in a number of service industries are also affected. These include caterers, garage mechanic and hairdressers.

It is now becoming accepted that skin and respiratory sensitisation occur by different immunological mechanisms [56, 19, 148, 17, 20, 149]. However there are a number of chemicals (perhaps more than 20) that have been reported as causing both asthma and dermatitis. These include the aldehydes (glutaraldehyde [150, 151, 152]) and some isocyanates (for example hexamethylene di-isocyanate [153, 154], toluene di-isocyanate [153] and diphenylmethane di-isocyanate [155, 156, 153]). Interestingly, no acid anhydrides appear to have been reported to cause skin sensitisation however some of the derivative acids such as hexahydrophthalic acid have [157].

1.4 Structure-Activity Relationship Studies

Overview

Structure-activity relationships (SAR's) are associations between a chemical's structure and its activity, whether the activity is a physical, chemical or biological property. To perform a SAR study one needs to be able to represent structure and activity in a manner in which qualitative or preferably quantitative descriptions can be made of the relationship between the two.

Chemical Structure Representation

A commonly held visualisation of chemical structure (at least amongst non-chemists) is that of the classical Lewis structure 'graph' two-dimensional diagrams. Lewis theory assumed that atoms bond by sharing electrons thus completing their shells without ionisation. Graph representation is so commonplace that there is a tendency for non-chemists to think of molecules solely in terms of such structural representations. This empirically drawn up representation of chemical structure is limited. The structure of benzene must be represented by at least two resonance structures using this format even though measurements suggest all the carbon-carbon bond lengths in benzene are equal. In reality chemical structure is far more complex. The exact forces that hold a collection of atoms within the relatively stable state seen in molecules are not well understood. These forces mean that molecules are dynamic. Chemical structures are not rigid.

Chemical structures can be represented in a number of increasingly descriptive ways. The methods of description have improved as larger, more complex chemicals have been described. Although chemicals can be described in empirical formula, such a description is too simplistic to use unless the molecules are very small. Chemical descriptions must be sufficient at least to describe the topological relationship between the atoms of a molecule. The topology of a molecule is the description of which atoms are connected to which other atoms. In addition it is sometimes important to specify the topography of a molecule. Topography refers to the positional co-ordinates in three dimensional space of each atom in a molecule. Two chemical structures may be topologically identical but topographically different (because for example they are stereo-isomers). It may also be necessary to a study that no two different chemicals have the same description. The definition of 'different' may depend on the context of the chemical study. In this study topologically equivalent compounds are considered identical even if they

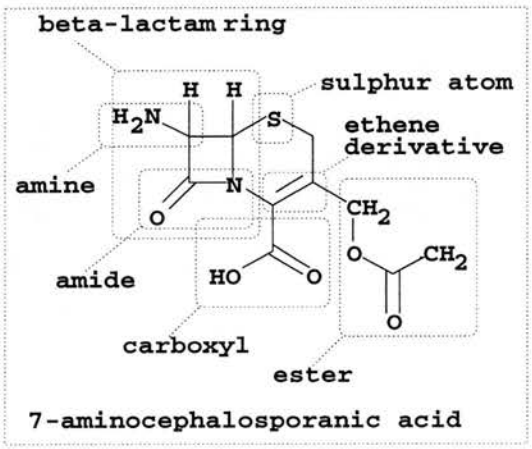


Figure 1.15: Chemicals can be described by their substructure fragment content.

differed stereo-chemically.

The storage and retrieval of chemical structure is well reviewed elsewhere [158, 2, 159]. Storage is a compromise between accurate representation and fast retrieval by querying. Accurate representation involves having tables of all connections of all atoms. One of the many accepted chemical representation formats is the 'molfile' format [159] from MDL[1] and it is this format that has been used in this study to store structure data. There is a complication with molfiles in that they do not necessarily include hydrogens [159]. The hydrogens are considered implicit (although they may sometimes be explicitly represented).

The approach used to describe structure in this work is similar to that described as atom/bond-centred by Bawden [158]. Briefly, molecules are described by a string of numbers representing the number of occurrences of particular substructure-fragments. These fragments may overlap (see Figure 1.15).

[1]See URL:

`http://www.mdli.co.uk/`

Predictive Systems

The fundamental assumption that underlies most attempts to predict properties from chemical structural data is that structurally similar molecules will exhibit similar properties [2]. Indeed it has been argued that properties of chemicals are ultimately a function of their structure:

> The structural formula of an organic compound, in principle, contains coded within it all the information which predetermines the chemical, biological and physical properties of that compound. [160]

Essentially, structure activity relationships are made by forming mathematical descriptions of the relationship between structure[1] and activity. Given such a mathematical description one can apply the same relationship to a chemical whose structure is known but whose activity is not. Hence mathematical description gives a prediction of activity based on prior observation.

The accuracy of prediction is based on the quality of the mathematical description of the relationship between structure and activity, both in terms of the validity of the mathematical expression and the reliability of the data. If chemical structures or activities are wrong the model will fail. Equally, if the mathematical expression does not attempt to exclude coincidental structure-activity relationships then the model will only be relevant to the original learning set of structures and activities. In short, each constant and coefficient in a predictive expression must be justified in terms of statistical validity.

The type of prediction produced by a predictive model may take one of several forms. Predictions can be true or false (qualitative); categorical (qualitative); or quantitative. The prediction may take the form of a

[1] Or a property clearly derived from the structure, such as molecular mass

probability of an outcome. So, for example, if predicting the weather a model might predict whether precipitation will occur; whether it will be rain, hail, sleet or snow; what the probability of rain is; or how many inches will fall.

This study of chemical sensitizing potential forms a prediction that can be taken as a binary value or as a probability. In essence, it addresses the question "is this compound a hazard" and the resulting prediction can either be taken as yes or no, or perhaps more usefully the probability that it is a hazard. It is *not* a measure of how hazardous a compound is.

Measuring Predictive Power

It is important to know how good a model is at prediction. Given that there are imperfections in every model it is useful to quantify how good a model is. The residuals - the difference between observed and predicted values - provide a means of doing this. With binary or classification prediction (in which the prediction will assign to a particular category) there can be measures of specificity and sensitivity (see p77). Sensitivity describes how well a model correctly assigns those values which belong in a category. Specificity describes how well the model excludes values from a category in which they do not belong. It is worth noting that sensitivity and specificity for the same predictive test are reciprocally related.

Validating Predictive Models

If a predictive model is to be of use it should be validated . A measure of how frequently it produces the correct result is needed. In order to achieve this a validation data set that is independent of the learning data set is required. A model than can then be tested on a previously

unseen data set to ascertain whether its predictive performance is of general use rather than merely specific to its learning data set. It is measurements such as predictive power, sensitivity and specificity calculated for a validation data set which best indicate whether a model is 'good'.

1.5 Aims

The aims of this thesis can be described in three parts: firstly, the collection of data; secondly, the use of that data to describe (and predict) occupational chemical sensitizing hazard using chemical structure; and thirdly the use of that relationship to stimulate novel mechanistic hypotheses. Throughout, the emphasis will be placed on finding a practical method of assessing occupational chemical sensitizing hazard.

Chapter 2

Methods

2.1 Overview

The study design consisted of three major stages: data collection; data analysis; and data validation. The first stage, data collection, involved the compilation of a database of control and asthmagenic chemicals. The second stage, the data analysis, involved the creation of a structure-activity relationship for these chemicals in order to model and predict their respiratory sensitizing ability. The third stage, the data validation, involved the testing of the predictive models resulting from the second stage with a previously unseen (validation) set of controls and asthmagens.

The study was not however restricted to the difference between asthmagens and control chemicals. From the outset, an additional set of chemicals known to cause skin sensitisation was to be compiled and this too compared with the set of asthmagens. It should stated however that a comparison between the skin sensitizers and the controls did *not* form part of the study design.

2.2 Data Collection

General Criteria

Searches of the medical literature up to the end of December 1994 were carried out using the MedLine Database[1]. Published cases of chemically induced OA and allergic contact dermatitis were assessed for inclusion in this study as the 'active' compounds. The disease had to be occupational in origin. A further restriction was to limit the papers to those describing human results using the MedLine option *limit to human* (for example, see option 10, Table 2.3). 'Control' compounds were selected from compounds in use which are known to be hazardous. The chemical identity (speciation) of all substances used had to be clear (in terms of a definitive chemical structure) . Organometallic complexes with transition metals where excluded but organic salts formed with other metal ions were accepted. There was no restriction on the allowable elements in the compound, however for the purposes of this study 'organic' was defined as 'carbon-containing'. (Therefore compounds that did not contain at least one carbon were rejected.) All compounds selected were of molecular mass less than 1000.

Asthmagens

Since the use of the MedLine 'keyword' system is not sensitive, merely specific, several textword based searches were performed in which words synonymous with the general topics of occupational disease, sensitisation and respiratory disease were used (see Table 2.1) . The initial asthmagen searches preceded the search for contact sensitizers but the later searches benefited from the lessons learnt from the skin search validation (see p55). All search results were added to a reference

[1]This is provided by CHEST Ovid Biomedical Service. See `http://biomed.niss.ac.uk/` for details.

occupational	asthma	sensitisation
work-related	rhinitis	sensitising/sensitizing
industrial	wheeze	allergy
industry	respiratory symptoms	allergic
manufacture	pulmonary oedema	hyper-sensitivity
manufacturing	anaphylactic shock	

Table 2.1: Asthma keywords used in multiple MedLine searches.

database (the package *Reference Manager* was used). The reference database software readily enabled checks to ensure papers were only entered once. Approximately 2000 references resulted which were reviewed by title and abstract and were potentially relevant by appraising the full paper to identify reports of respiratory sensitizers.

The references identifying respiratory sensitizers were studied in depth and the potentially relevant available epidemiological data tabulated[1]. Data obtained included reported prevalence, type of response, evidence of IgE involvement, latency and persistence of disease. These data were tabulated by categorising (See Table 2.2). Many of the fields remained blank or reported 'no data available' to indicate the fact that such data were sought. A record of the relevant references and any additional miscellaneous notes (e.g. for glutaraldehyde the field states 'Highly toxic, corrosive, protein cross-linking agent') were also made.

Active chemicals were defined as substances for which a physician had in a peer reviewed report clearly diagnosed OA arising following a latent period of exposure.

[1]This aspect of the study could be viewed as prospective since during the recording stage it was not clear which epidemiological data would prove useful. As a consequence the recording was systematic and unbiased.

Table 2.2: Format of the clinical data coding protocol

Category	Coding Options
Absolute prevalence	0. No known cases
	1. A single case report
	2.Up to 3 cases
	3. Up to 10 cases
	4. Up to 33 cases
	5. Up to 100 cases
	6. Up to 333 cases
	7. More than 333 cases
Percentage prevalence	0. No data available
	1. Less than 1%.
	2. In the range 2-4%
	3. In the range 5-9%
	4. In the range 10-24%
	5. In the range 25-49%
	6. In the range 50-74%
	7. In the range 75-100%
Sensitising Concentration	0. No data available
	1. Less than 0.0001 μMoles/l
	2. Less than 0.001 μMoles/l
	3. Less than 0.01 μMoles/l
	4. Less than 0.1 μMoles/l
	5. Less than 1. μMoles/l
	6. Less than 10 μMoles/l
	7. Less than 100 μMoles/l
	8. Less than 1000 μMoles/l
	9. More than 1000 μMoles/l
Type of response	0. No data available
	1. Immediate
	2. Late

	3. Immediate or late
	4. Dual
	5. Immediate or dual
	6. Late or dual
	7. Early, late or dual
IgE testing	0. No data available
	1. No evidence found.
	2. Elevated total IgE levels
	3. Specific IgE found
Basis for diagnosis	0. No data available.
	1. Workplace observation
	2. Respiratory challenge.
Percentage atopic	0. No data available
	1. Less than 1%
	2. Less than 5%
	3. Less than 10%
	4. Less than 25%
	5. Less than 50%
	6. Less than 75%
	7. More than 75%
Latency (median)	0. No data available
	1. Up to 1 month
	2. Up to 6 months
	3. Up to 1 year
	4. Up to 2 years
	5. Up to 3 years
	6. Up to 5 years
	7. Up to 10 years

Controls

Similar chemical criteria were employed for the inclusion of control compounds. The control chemicals were selected from the Tables 1 and 2 of the Health & Safety Executives EH40/94 Occupational Exposure Limits 1994 document which detail occupational exposure indices [161]. All the listed organic compounds (excepting those which were already included in the study as active) and for which a structure could be identified were used. If a limit was set for all structural isomers of a compound, then all those isomers were included.

Contact Allergens

The criteria for selection of the occupational skin sensitizers closely mirrored those for the asthmagens and the controls. The storage of the compounds was done using the IsisBase package[1]. The significant difference between the collection of the asthmagens and the skin sensitizers was that for the latter an upper limit of 200 compounds was set. Whilst considerably in excess of 200 skin sensitizers are known [162], not all of these are cited as occupational skin sensitizers. The specific criteria needed for a compound to be included as a skin sensitizer were: the compound contained carbon; the molecular weight was less than 1000; the compound caused *allergic contact* dermatitis; and that the compound caused the disease through occupational use. A systematic search through MedLine was performed using the MedLine search criteria detailed in Table 2.3. The results of this search were then examined in order to identify possible candidate chemicals by evaluating the title and abstracts. Compounds for which there was doubt about whether they fulfilled the chemical structure criteria were at this stage

[1]A chemical structure database available form MDL Inc., see `http://www.mdli.co.uk/` for company details.

MedLine Search
1. exp "dermatitis, allergic contact"/
2. industrial.tw
3. work.tw
4. occupational.tw
5. "dermatitis, occupational"/
6. 2 or 3 or 4 or 5
7. allergic.tw and 4
8. 1 and 6
9. 7 or 8
10. limit 9 to human

Table 2.3: Initial MedLine search strings for the identification of skin sensitizers.

still included. For the years 1966-1975 additional textwords (to recognise *allergic contact dermatitis*) had to be used as MedLine's subject headings were not extensively used over that period (i.e., the first option in Table 2.3 was ineffective).

A validation search was then performed using a exhaustive trawl of one year's publications in perhaps the most authorative journal of dermatitis published in English - Contact Dermatitis. The year 1992 was selected semi-randomly - a fairly recent year was needed because the local source of Contact Dermatitis did not have the full set of years for the journal. Using Contact Dermatitis, Volume 26, 1992, all possible candidate compounds were identified. These were compared with compounds previously identified from the MedLine search as originating from Contact Dermatitis, Volume 26, 1992. A number of compounds not previously identified arose in this validation search indicating the shortcomings of the original searching terms. Using the data from the validation search, further search criteria were applied in addition to those in Table 2.3. These were entered as textwords and are described in Table 2.4. They fell into three classes - synonyms of *occupational*, synonyms of dermatitis and synonyms of *allergic*. In particular the spelling of sensitisation with a *'z'* or *'s'* had to be accounted for (see

'occupational'	'dermatitis'	'allergic'
work	eczema	sensitised
industrial		sensitisation
industry		sensitized
manufacture		sensitization
manufacturing		hyper-sensitivity

Table 2.4: Additional criteria added to the search for skin sensitizers.

Table 2.4).

The effectiveness of the first literature search for contact sensitizers was barely 20%. Of 28 references considered relevant in Volume 26 of Contact Dermatitis 1992 only 6 had been found in the initial search (see Table 2.3). From this it was evident which keywords were missing from the search (see Table 2.4).

Having identified in excess of 200 potential candidate compounds the definitive chemical structures were obtained using a number of sources, principally the Beilstein Crossfire service (then available on trial from Daresbury Laboratory but now part of MIDAS). Other sources of structures were the Merck index, the Aldrich catalogue and the Chemical Abstracts. After the exact structure was determined those compounds which no longer satisfied the study chemical structure criteria were discarded.

The Validation Chemicals

A validation set of chemicals was prepared by Dr Raymond Agius[1]. Control chemicals from later editions of the EH40 Occupational Exposure Limits documentation were used. Asthmagens were selected from cases of chemically induced OA published from 1995 onwards. Suspected asthmagens were kindly made available from those reported to

[1]Dr Raymond Agius, Department of Public Health Sciences, The University of Edinburgh.

	Controls	Asthmagens	Skin sensitizers
Learning set	≥ 100	~100	200
Validation set	?	?	N.A.

Table 2.5: Summary of the Expected Content of the Data Sets.

the SWORD[1] scheme [27, 163]. All the validation chemicals were systematically checked to ensure that the structures had not been present in the learning data sets under different names.

Summary of the Data Sets

The data sets to be used were expected to contain approximately the number of chemicals described in Table 2.5. There is some overlap of some of the data sets (see Figure 2.1), indicating that some compounds were common to both data sets. The relevance of any overlap depended upon which data sets were being compared (see Figure 2.2). So for the analysis of the differences between skin and respiratory sensitizers, the components in the overlap 'A' of Figure 2.1 were necessarily excluded. These compounds (in overlap 'A', Figure 2.1) could however be included in the analysis to distinguish between asthmagens and non-asthmagens, as indeed could the compounds in overlap 'B'. The stated definitions for asthmagens (see p50) and controls (see p54) are independent of the definition given for skin sensitizers (see p54), hence the comparison between the two is unaffected by the skin sensitising status of the compounds. Finally, the compounds in overlap 'C' (Figure 2.1) which were controls in the learning set could be 'promoted' to the set of validation asthmagens (which were selected later).

[1]Surveillance of Work-related and Occupational Respiratory Disease, a scheme introduced in 1992 to allow chest physicians to report occupational groups and agents with a high risk of respiratory disease (including asthma).

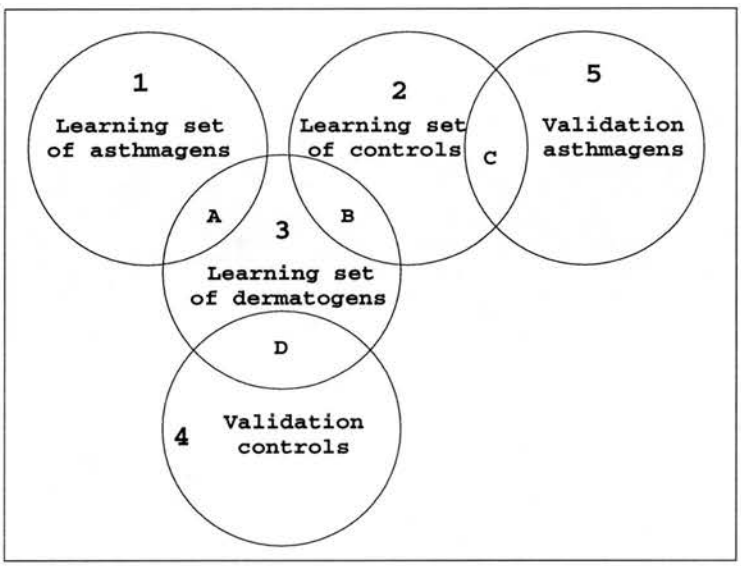


Figure 2.1: Summary of Data Set Overlap.

2.3 Chemical Structure Data

Storing Chemical Structures

Chemical structures were stored as MDL *.mol files [159] (see Figures 2.4 and 2.3) inside the MDL IsisBase package (IsisBase 1.2.1, MDL14, [159]). The compounds in the database were analysed using a modified case-control approach [164]. The case-control study methodology was adapted to use chemical entities rather than subjects as the study population (see p18). The 'cases' were reported causative agents of OA, whereas the 'controls' were matched hazardous occupational agents which were not reported causes of OA. The 'exposures' were the selected chemical substructure fragments contained within the population.

The chemical structure drawing package IsisDraw[1] was used to draw

[1]Available free for non-commercial and academic users from MDL - see URL http://www.mli.co.uk/ for details.

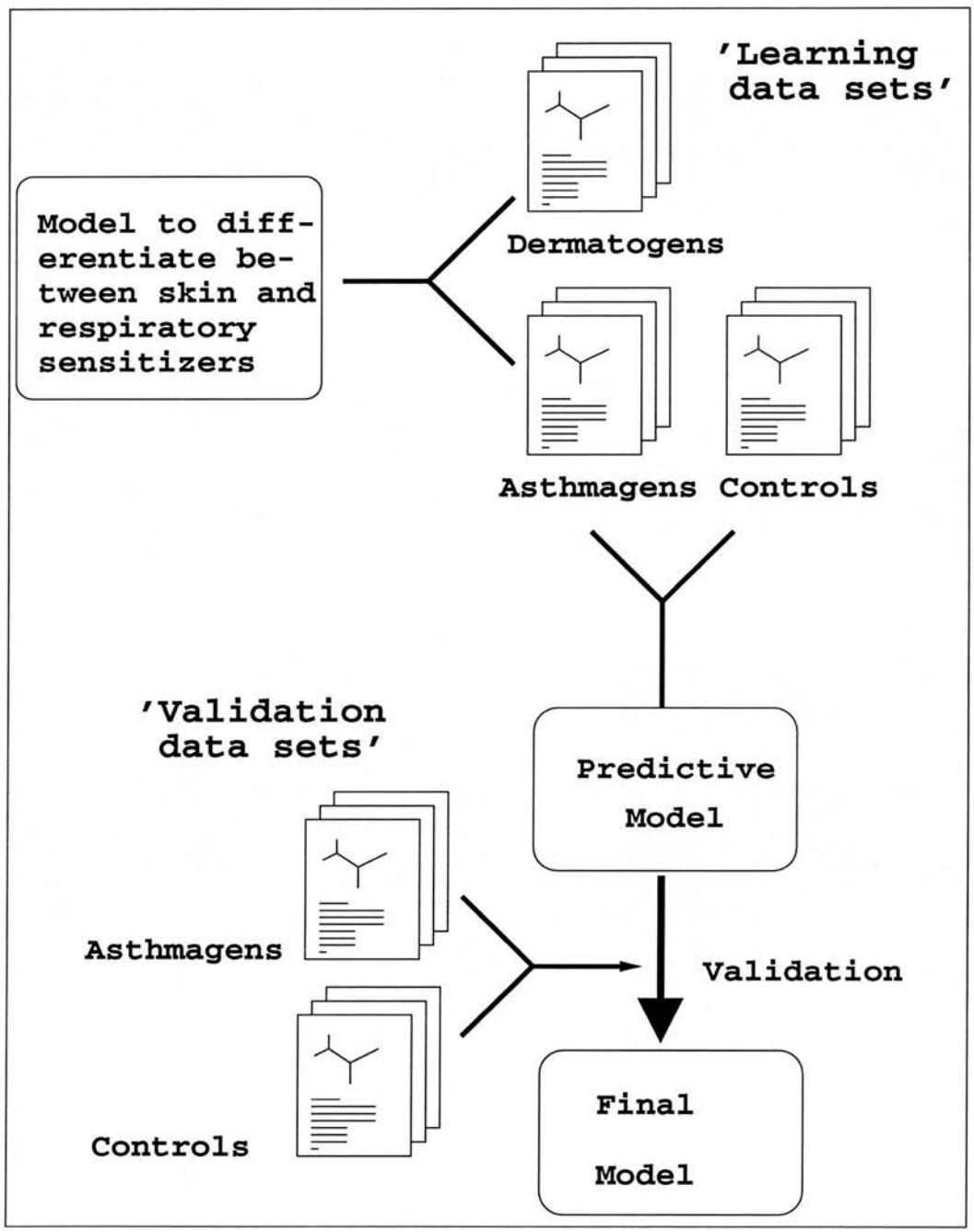


Figure 2.2: Overview Showing Which Data Set Pair Comparisons Were Made.

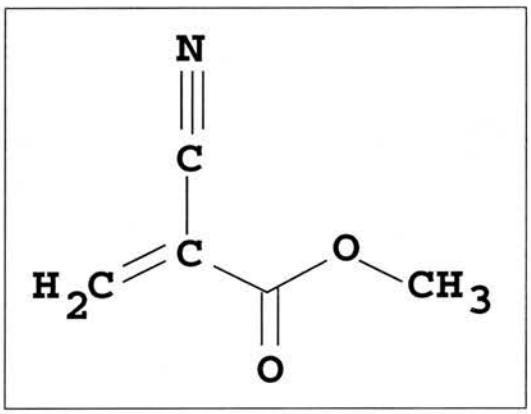


Figure 2.3: Methyl-2- cyanoacrylate as represented by the MDL *.mol molfile in Figure 2.4.

in the structures. The molfile format is a chemical structure connection table with a *header*, a *counts line*, an *atom block* and a *bond block*.

The *header* consists of up to 3 lines: a comment or molecule name line; a line detailing the type of application which generated the table, the date and time of its creation; and a further comment or description line that even if empty must be present.

These lines are followed by the *counts line* which gives the number of atoms and bonds (plus several additional flags) in the molecule.

There then follows the *atom block*, containing the x, y, and z co-ordinates of the atom and the element symbol for the atom. As this study used only topological (rather than topographical) structural information the positional coordinates were not used.

Finally, for the *bond block*, the first two numbers of each line indicates the numbers of the respective atoms from the atom list, and the third number indicates the bond type between these two atoms. (A value of 1 indicates a single bond connection, 2 a double bond and 3 a triple bond.) The full MDL molfile specification does describe further "flags" but these were not used [159].

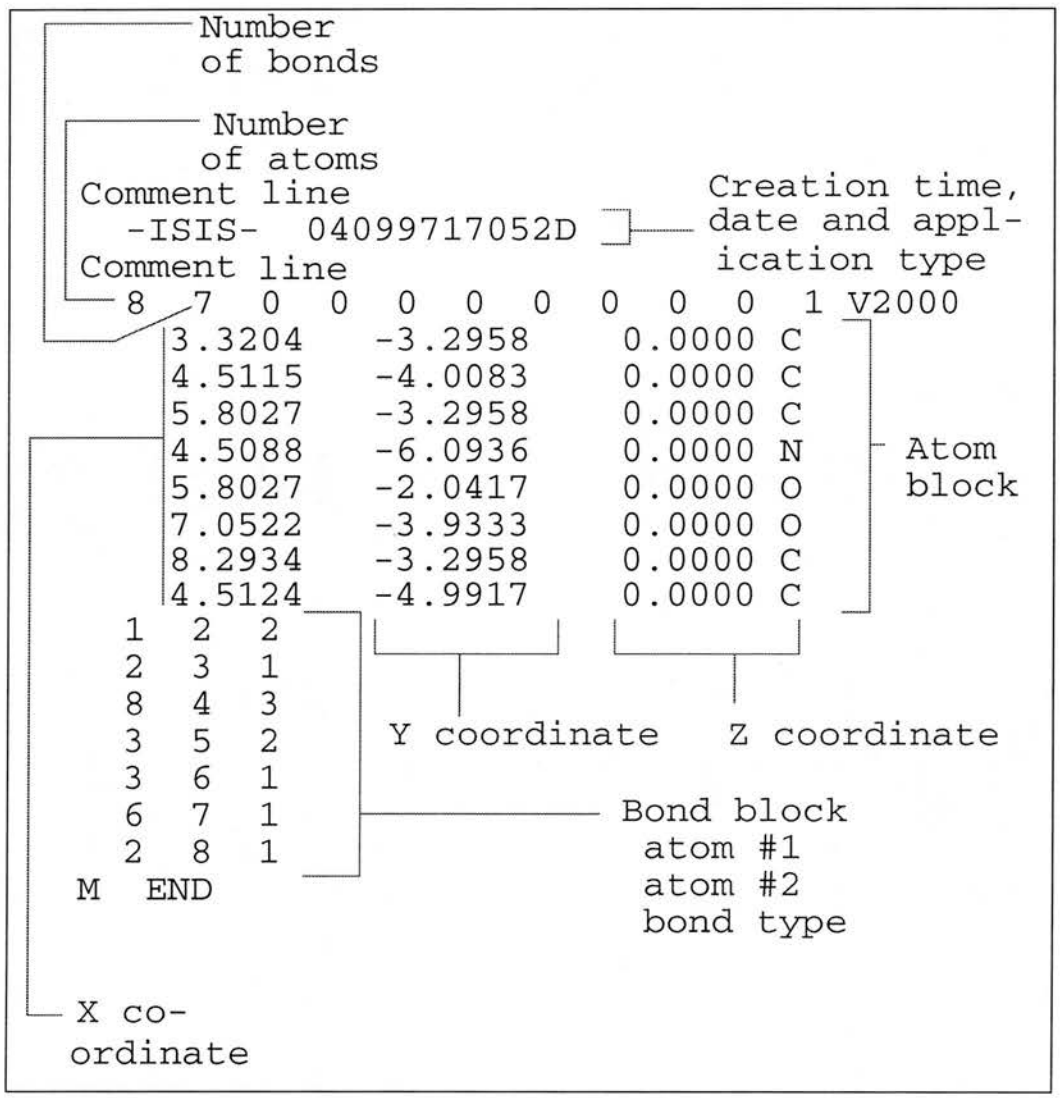


Figure 2.4: Example of an MDL Molfile.
This figure shows a connection table for methyl-cyanoacrylate (see 2.3) Additional redundant flags to lines of the *atom* and *bond blocks* have been omitted for clarity (see 2.5 for the full version).

C	N	...	Ar_N	XCCX	category	MOLMASS	MOLNAME
9	0	...	0	0	0	108.10	3-Methylstyrene
6	4	...	0	4	1	128.09	Hexamethylene_tetramine

Table 2.6: Structure represented by fragment occurrence frequencies.

Large numbers of structures and their property data could be collectively stored using an identical connection table format in MDL `*.sdf` files [159], an ASCII file format (see Figure 2.5) available when using IsisBase. Chemical and clinical property could also be stored in these files. The data for each molecule follows the connection table. Separate molecule entries are demarcated by the `$$$$` flag on a line by itself.

Describing Structures: Fragments

Throughout this study the term fragment is used to refer to the type of chemical structure descriptor used. Fragments were selected from the represented atom types and known chemical groups observed within the set of chemical structures studied. The selection of fragments was partly subjective. The fragments selected tended to be small and fairly commonly occurring within the dataset of chemicals. Fragments were not necessarily distinct chemical structures. The resulting tabulated fragment occurrence frequencies for each of the studied chemical structures provided a numerical structure description (see Table 2.6)[1].

Two types of fragment were used in this study (see Figure 2.6). The first may be represented in a standard connection table[2] - however it is not necessarily a distinct structure since unfilled valencies do not imply hydrogen atoms. (Unfilled valence could be matched to any atom type

[1]The fragment content description may be the same for two topologically different molecules. For example the various isomers of xylene (that is dimethyl benzene) are topologically distinct but have the same fragment description.

[2]The 141 Type 1 fragments used may be obtained from the file : `[CDROM_DRIVE_LETTER]:\data\sdf\newfrag.sdf` on the CD referred to in Appendix B.

```
 -ISIS-  04269517002D

  8  7  0  0  0  0  0  0  0  0  1 V2000
    3.3204   -3.2958    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
    4.5115   -4.0083    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
    5.8027   -3.2958    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
    4.5088   -6.0936    0.0000 N   0  0  0  0  0  0  0  0  0  0  0  0
    5.8027   -2.0417    0.0000 O   0  0  0  0  0  0  0  0  0  0  0  0
    7.0522   -3.9333    0.0000 O   0  0  0  0  0  0  0  0  0  0  0  0
    8.2934   -3.2958    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
    4.5124   -4.9917    0.0000 C   0  0  0  0  0  0  0  0  0  0  0  0
  1  2  2  0  0  0  0
  2  3  1  0  0  0  0
  8  4  3  0  0  0  0
  3  5  2  0  0  0  0
  3  6  1  0  0  0  0
  6  7  1  0  0  0  0
  2  8  1  0  0  0  0
M  END
>  <MOLNAME> (53)
Methylcyanoacrylate

>  <CAS#> (53)
137-05-3

>  <Merck#> (53)
5667

>  <category> (53)
1. Respiratory sensitiser.

$$$$
```

Figure 2.5: Format of a single record (for methyl cyanoacrylate) of an MDL `*.sdf` file.
Properties of the chemical can easily be read, for example the CAS number property (CAS#) for this compound is [137-05-3]. The `$$$$` line indicates the end of the record. This example only shows three property fields, in the full database there were many more. Note: In Figure 2.4 the additional flags in the atom and bond blocks shown here were omitted.

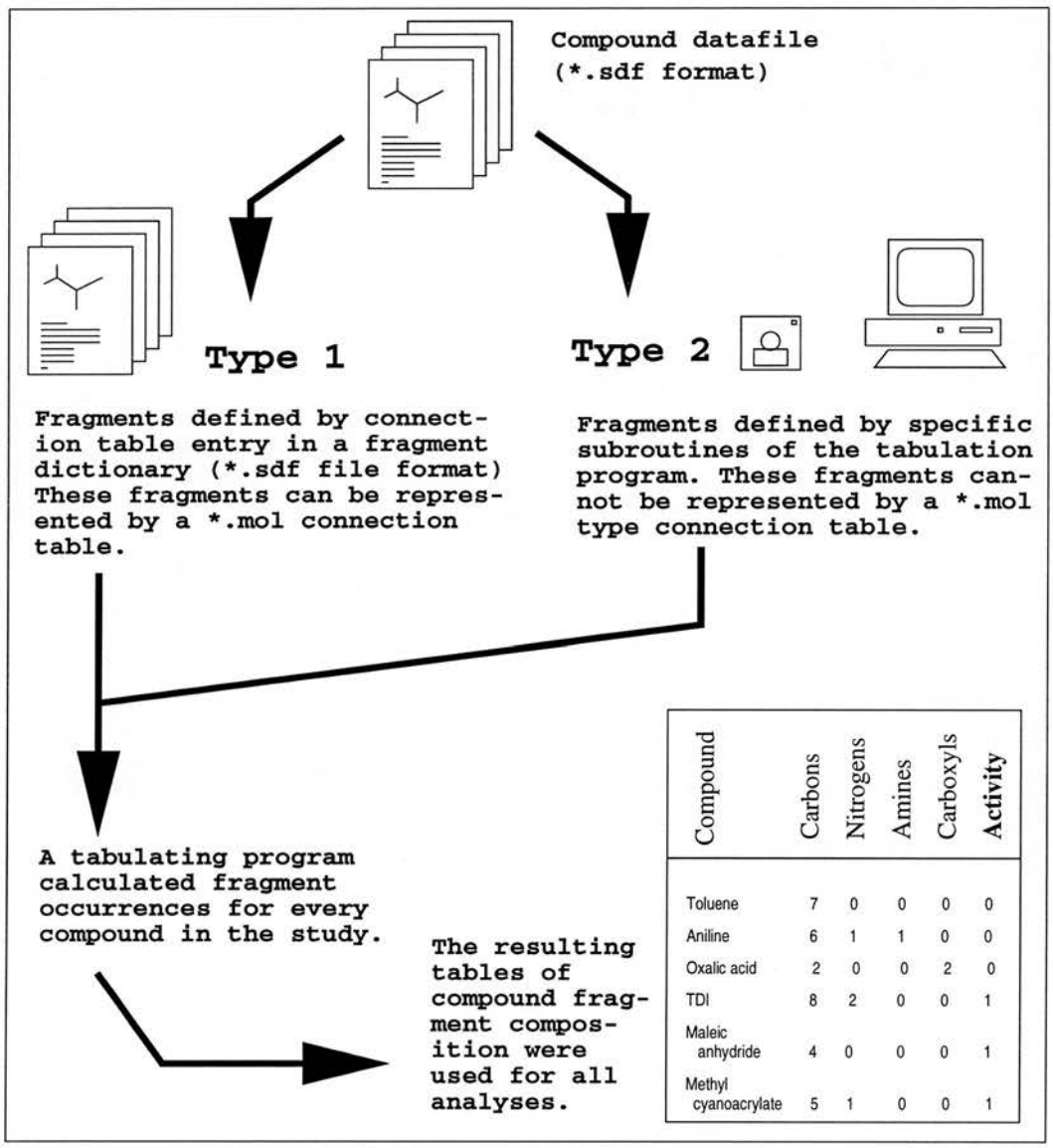
in the molecule under investigation.) The second type, (Type 2 Fragments) are characterised by having rules affecting the degree of variability tolerated for a bond or node (atom position) in a skeleton structure. These rules are hard coded in a computer program subroutine specific to that fragment. Twenty eight such subroutines were written (based on information derived from Type 1 fragment analyses)[1]. In addition a atom-finding routine was also included that calculated the number of occurrences of a specified atom (see Appendix B).

Both the Type 1 and the Type 2 fragments were chosen subjectively by consideration of the possible fragments available in the dataset and by the requirement that the fragment should occur sufficiently frequently so as to have the potential to be of statistical significance. Although Type 1 fragment selection was subjective a degree of method was employed to ensure that a moderately systematic approach was

To illustrate the difference between the two techniques it is appropriate to use an example. Consider the amine chemical group, a group which is centred around a nitrogen single bonded to at least one carbon. Amines can be 1°, 2° and 3° (depending on whether they have two, one or no hydrogens attached to the nitrogen respectively). If we wish to have a single fragment representing all possible amine types, 1°, 2° and 3° we have a problem with Type 1 fragments. We cannot specify a single connection table which will match 1°, 2° and 3° which will not also match several fragments which are not amines (see Figure 2.7)[2]. Figure 2.7 does not show an exhaustive list of compounds which would match the fragments. An additional consideration is the presence of atoms bonded to the carbon atom - for example an oxygen

[1]For the definitive description of these fragments refer to the program source code in the file
`[CDROM_DRIVE_LETTER]:\programs\sdf\newfrag.sdf` on the CD

[2]It may seem odd to want to combine the three amine types as in doing so one loses structural detail however it should be remembered that for a fragment to be of relevance it must occur frequently in the set of compounds. By combining certain similar groups, occurrence frequencies are more likely to approach values which will yield statistically significant results.



Compound	Carbons	Nitrogens	Amines	Carboxyls	Activity
Toluene	7	0	0	0	0
Aniline	6	1	1	0	0
Oxalic acid	2	0	0	2	0
TDI	8	2	0	0	1
Maleic anhydride	4	0	0	0	1
Methyl cyanoacrylate	5	1	0	0	1

Figure 2.6: Two Types of Fragment (Types 1 and 2) Were Used

double bonded to this carbon would make the group an amide rather than an amine group. Type 2 fragments overcome this problem because one can specify that the atom at a particular node must be of element *A or B but not C*. The more flexible rule based criteria of Type 2 fragments can also be applied to bonds - for example one can have a fragment that matches all occurrences of carbon bonded to nitrogen irrespective of whether it is by a single, double or triple bond. It may appear that Type 2 fragments are far better but they do have a major drawback. Each fragment must be hard coded in a subroutine, which requires writing and testing.

In contrast Type 1 fragments can easily be added to an *.sdf dictionary[1] of fragment structures merely by drawing them with IsisDraw and copying them into IsisBase. The tabulating program then reads a connection table for the fragment and calculates the occurrence frequency in each chemical structure studied.

The two types of fragments were employed effectively as follows: Type 1 fragments were used to screen a large dictionary of chemical substructures (up to 130) to identify those which were disproportionately distributed between the active and control compounds studied; then, having identified the most relevant Type 1 fragments, these were refined into Type 2 fragments.

The creation of fragment occurrence tables for compounds was performed using two computer programs (one for each fragment type) written in the programming language 'C'. This was necessary both to reduce the time required to compile the tables and also to ensure consistency of fragment matching. In theory the programs should be portable to any machine with a suitable compiler (although users of MS-DOS / Windows 3.x may experience difficulties with memory allocation).

[1]A fragment dictionary is simply a collection of fragments against which matches could be made.

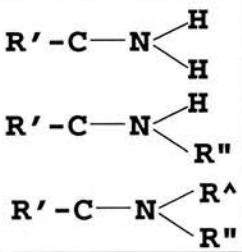
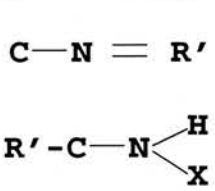
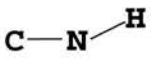
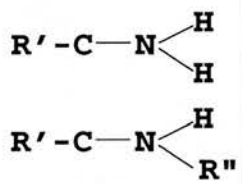
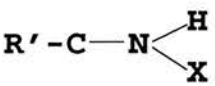
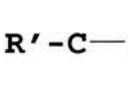
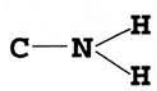
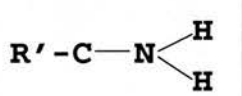
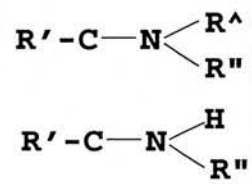
Fragment	Matches	Also Matches	Does not Match
			
			
			

Figure 2.7: How fragments match.

2.4 Data Analysis

The analytical methods used for interpreting the data included: clustering techniques based on graph theory; adapted odds ratios which in this text are termed *hazard odds ratios* (HOR's); and the use of applied logistic regression modelling. All these methods depend upon the *a priori* assumption that structurally similar molecules will exhibit comparable properties [2] consequently each method represents a test of the null hypothesis (see p15).

Clustering

Clustering is a technique by which "entities" (in this case chemical structures) are assigned to groups according to similarity. To achieve this the entities must be compared by means of "descriptors" (for example in this study chemical substructure fragments are used to describe the chemicals).

For each chemical studied a point in N-dimensional space can be assigned. N is the number of descriptors (substructure fragments) used to describe each compound. So for n^{th} descriptor (fragment), the position along the n^{th} axis is given by the value of that descriptor. In the case of a substructure fragment, the position along the n^{th} axis equates to the number of times the fragment has occurred in the compound. The result is that each compound can be described by a grid reference like a position on a map, except that there are generally more than 2 dimensions!

The distance between entities (compounds) is a measure of their similarity (or dissimilarity). The distance can be easily calculated using classical Euclidean geometry: the distance $d_{i,j}$ between the two objects

(i and j) is given by Equation 2.1 where N is the total number of descriptors and i_n and j_n are values for n^{th} descriptor of entities i and j respectively [165].

$$d_{i,j} = \sqrt{\sum_{n=1}^{N}(i_n - j_n)^2} \tag{2.1}$$

It is possible for the descriptors to take the form of binary or continuous variables. The similarity measures predominantly used in this thesis uses n *continuous* variables. Where not explicitly mentioned continuous variables should be assumed. For some of the clustering analyses (those based on presence or absence of substructure fragments) the n variables took a binary form. These latter similarity measures will be identified as such by explicitly indicating that *binary* variables were used.

Three clustering methods are described. The first two - the Ward and the Jarvis-Patrick methods - are only briefly detailed as the clustering done using these methods was performed on a preliminary dataset by Dr. Geoff Downs[1]. The third method - the Guenoche method - was also used by Dr Downs on the preliminary dataset and this appeared to produce the most reliable results. This was in agreement with an earlier study [166]. Consequently, following discussion with Dr Downs, the Guenoche method was implemented for further clustering studies.

The Ward Method of Clustering

The Ward Method of Clustering [166] involves identifying all the reciprocal nearest neighbour pairs (i.e. for pairs x and y, x is y's nearest neighbour and y is x's nearest neighbour). The method then merges

[1]Barnard Chemical Information Ltd.

these pairs into a hybrid pair. It was only used for the preliminary study by Dr Downs.

The Jarvis-Patrick Method of Clustering

The Jarvis-Patrick Method of Clustering [167, 166] is a non-parametric technique based on shared nearest neighbours. Two points are assigned to the same cluster if they share a set threshold number of nearest neighbours. The two points must also contain each other in that nearest neighbour list. It too was only used for the preliminary study by Dr Downs.

The Guenoche Method of Clustering

The Guenoche Method of Clustering is an example of divisive hierarchical clustering. It uses a parametric method. It depends upon a table of the $2N$ differences between all pairs of N compounds. The algorithm is:

1. Calculate the distances between all possible pairs of points.
2. Select the most distant pair of points x and y (which are in the same cluster).
3. Assign point y to a new cluster.
4. For all remaining points in the original cluster identify whether they are nearer to x or y.
5. Assign points nearest to y to the new cluster (cluster containing point y)[1].
6. Repeat steps 1-5 until desired number of clusters is reached.

A C program was written to perform cluster analysis (see Appendix B). There was no obvious choice for the number of clusters to generate so steps of five were initially used. Later cluster sizes of 1, 2 3, 10, 30, 100 and 300 were used in order to cover a wide range of cluster sizes.

Clustering does not readily lend itself to prediction. The method used to evaluate its potential as a predictive method was as follows: the full set of compounds was clustered into n clusters; the clusters were then categorised as predictive of asthmagens or controls on the basis of the relative proportions of each in the cluster (see Equation 2.2). For example, if a cluster had more asthmagens in it than would be expected given the ratios of asthmagens to controls in the full dataset, that cluster was deemed predictive of an asthma hazard. A 2x2 table of prediction activity versus observed (true) activity was then created and a kappa[1] value calculated (see Equation 2.5).

$$EXPECTED_{asthmagens} = \frac{TOTAL_{asthmagens}}{(TOTAL_{asthmagens+TOTAL_{controls}})} \times TOTAL_{cluster} \tag{2.2}$$

The source code of the program to Guenoche cluster and calculate a kappa value is included on the accompanying CD-ROM (see Appendix B).

χ^2 Testing of Clusters

The χ^2(chi-squared) test was applied as a crude measure of the predictive performance of clustering. As with other applications of two by 'n' tables, χ^2 can be calculated to see if the independent variable (in this

[1]It is possible that none of the remaining points are closer to y than x. In such a case y would be in a cluster by itself.

[1]Note a true kappa value as it is not calculated with a validation dataset, rather a kappa based on the learning dataset.

case the fragment) is associated the dependent variable (reported sensitisation hazard). (It should be noted that an *a priori* assumption has been made. The χ^2 value is merely a measure of association and does not indicate which variable is the dependent or independent.) The χ^2 test involves calculation of the expected cell frequencies in the two by two table. If E_a, E_b, E_c and E_d refer to the expected cell values for a, b, c and d respectively then:

$$E_a = \frac{(a+b) \times (a+c)}{(a+b+c+d)}$$

$$E_b = \frac{(a+b) \times (b+d)}{(a+b+c+d)}$$

$$Ec = \frac{(a+c) \times (c+d)}{(a+b+c+d)}$$

$$E_d = \frac{(b+d) \times (c+d)}{(a+b+c+d)}$$

The χ^2 value is given by the Equation 2.3.

$$\chi^2 = \frac{(Observed - Expected)^2}{Expected} = \frac{(a-E_a)^2}{E_a} + \frac{(b-E_b)^2}{E_b} + \frac{(c-E_c)^2}{E_c} + \frac{(d-E_d)^2}{E_d} \tag{2.3}$$

For a two by two table there is just one degree of freedom (rather than three) since the expected values are calculated from the marginal totals resulting in the loss of two degrees of freedom. The computer package Epi-Info[1] was used to calculate both the 2x2 $\chi^2$and the χ^2 for trend.

[1]Public domain software available on the cdrom (see Appendix B)

Hazard Odds Ratios

An Odds Ratio for the Hazard (or HOR) was used to characterise substructure fragments. This was the ratio of the two odds - the odds of a chemical being a reported causative agent given the presence of the substructure fragment; compared to the odds of a chemical being a case given the absence of the substructure fragment in question.

This is given by the equation:

$$HOR = \frac{a}{b} / \frac{c}{d} \tag{2.4}$$

where a, b, c and d are described in Table 2.7.

The calculation of HOR's was done using the Epi-Info statistical computer program package available from the Centers for Disease Control, Atlanta, Georgia, USA. The Exact Method was used when low expected occurrence frequencies caused the Cornfield confidence limits to be unreliable. HOR's were calculated with upper and lower 95% confidence intervals (C.I.s). A HOR was deemed significant if the range between the lower and upper 95% confidence intervals excluded the value 1.00. The range of value an odds ratio can take is between 0 and infinity.

Confidence intervals for odds ratios can be calculated by a number of methods. The simplest involve approximations of the probability distribution such as that used in the Taylor Series. Two methods of calculation of 95% C.I.'s were used.

HOR's by Occurrence Frequency χ^2 for Trend

Were HOR's are calculated for different occurrence frequencies these refer to the odds of a chemical being a reported causative agent given

and from the Epidemiology Program Office, Atlanta Georgia. See URL `http://www.cdc.gov/epo/epi/epiinfo.htm`

		Asthmagen?	
		y	n
Contains	y	a	b
Fragment	n	c	d

Table 2.7: 2x2 table for odds ratio calculation.

the presence n occurrences of the substructure fragment; compared to the odds of a chemical being a case given the absence of the substructure fragment in question. Where a χ^2 for trend is calculated, this was done using the Epi-Info package (see Appendix B).

Taylor Series

The approximate method uses the equation:

$$C.I. = HOR \times e^{1.96 \times \frac{1}{\sqrt{(\frac{1}{a}+\frac{1}{b}+\frac{1}{c}+\frac{1}{d})}}}$$

were 1.96 is the z value used for C.I.'s of 95%. This method is known as the Taylor Series. The Taylor Series method is inappropriate for data in which any of the expected cell values in the 2x2 table are small. It provides satisfactory results when *expected* cell values in the two by two table are not less than 5.

Fisher Exact Method

The need for an exact method of confidence interval (C.I.) measurement arises when cell values in the two by two table (Table2.7) are small such that expected values (see Table 2.8) are less than 5. In Table 2.8 if we consider the marginal totals (m_1, m_0, n_1, n_0) as fixed then for any change in E_a the values for the other three cells can be calculated. The statistical theory underlying exact testing is conditional on these fixed marginals. The calculations were performed by the Epi-Info statistical package.

		Asthmagen?		
		y	n	
Contains	y	Ea	Eb	n1
Fragment	n	Ec	Ed	n0
Marginal Totals		m1	m0	N

Table 2.8: Values in the Fisher Exact method of calculating confidence intervals.

		Observed	
		1	0
Predicted	1	a	b
	0	c	d

Table 2.9: 2x2 table for test agreement.

Test Agreement Statistics

Kappa values for the predictive value of clustering and logistic regression models were calculated to indicate the model agreement with observed data (See Equation 2.5) [168].

$$\kappa = \frac{\frac{a+d}{N} - \frac{E_a+E_d}{N}}{1 - \frac{E_a+E_d}{N}} \tag{2.5}$$

where

a = number of predicted cases that are true cases
d = number of predicted inactives that are true inactives
E_a = expected number of true cases
E_d = expected number of true inactives
N = total of all cases and all inactives

The values a and d along with their complementary expected values are derived from Table 2.9.

Kappa values below 0.4 indicate poor agreement; values between 0.4 and 0.5 indicate moderate agreement; a value between 0.5 and 0.6

indicate acceptable agreement; a value between 0.6 and 0.8 indicates good agreement; and values above 0.8 indicate excellent test agreement [168].

Logistic Regression

Logistic regression analysis is similar to ordinary linear regression, the principal difference being that the variables being modelled are discrete rather than continuous. In this study the modelled variable is sensitizing activity - true (1) or false (0). The independent variables (in this study the fragment occurrences) can be continuous. The resulting dependent variable (the prediction of activity) takes a continuous value between 0.0 and 1.0.

The logistic regression equation for each of the 1 to N fragments the logistic regression equation that gives the hazard index N for a given compound is:

$$y = \frac{e^{(\alpha+\beta_1 x_1+\beta_2 x_2+....+\beta_k x_k)}}{1+e^{(\alpha+\beta_1 x_1+\beta_2 x_2+....+\beta_k x_k)}} \tag{2.6}$$

Where α	=	a constant
β_1	=	the logistic regression coefficient for fragment 1
x_1	=	the number of occurrences in the molecule of fragment 1
β_2	=	the logistic regression coefficient for fragment 2
x_2	=	the number of occurrences in the molecule of fragment 2
etc...		
β_k	=	the logistic regression coefficient for fragment k
x_k	=	the number of occurrences in the molecule of fragment k

Logistic regression analysis was performed using the Statistical Pack-

age for Social Scientists (SPSS, from SPSS Inc.) package[1]. Data were entered in the package in the form of an ASCII file containing a line of information for each compound (See Table 2.6). Each line (or row) contained the chemical name, the activity (coded as 0 or 1 for control or active respectively) and a list of fragment occurrence frequencies. Although in excess of 130 substructure fragments were available the limit for SPSS was 60 and therefore a reduced set was used. The 60 chosen were selected after consideration of the HOR results.

A model was obtained using a backward stepwise method. The significance of the 'likelihood ratio' was used to determine which independent variables remained in the model. Variables were entered and removed from the model on the basis of an entry probability of 0.05 and a removal probability of 0.10. The Wald statistic was not used. Other model methods and parameters were tested but are not described.

The models were tested using a kappa test statistic based on the number of compounds with predicted hazard greater or less than a given threshold. The standard threshold was 0.5 although the effect of using other threshold levels on the kappa statistic was studied. An iterative process of improving the model by adjusting the fragments was used (see Figure 2.8)

Sensitivity and Specificity Sensitivity and specificity are measures used to evaluate the effectiveness of screening / diagnostic tests. Since the prediction of a hazard value for chemicals is in effect a screening test sensitivity and specificity are an appropriate additional measure of prediction model performance. Since the logistic regression approach yields a hazard index value (between 0 and 1) which is a probability of activity it can be useful to specify a threshold value (particularly if a test agreement statistic, kappa, is required). By adjusting the

[1]SPSS Inc, 444, N. Michigan Avenue, Chicago, Illinois 60611.

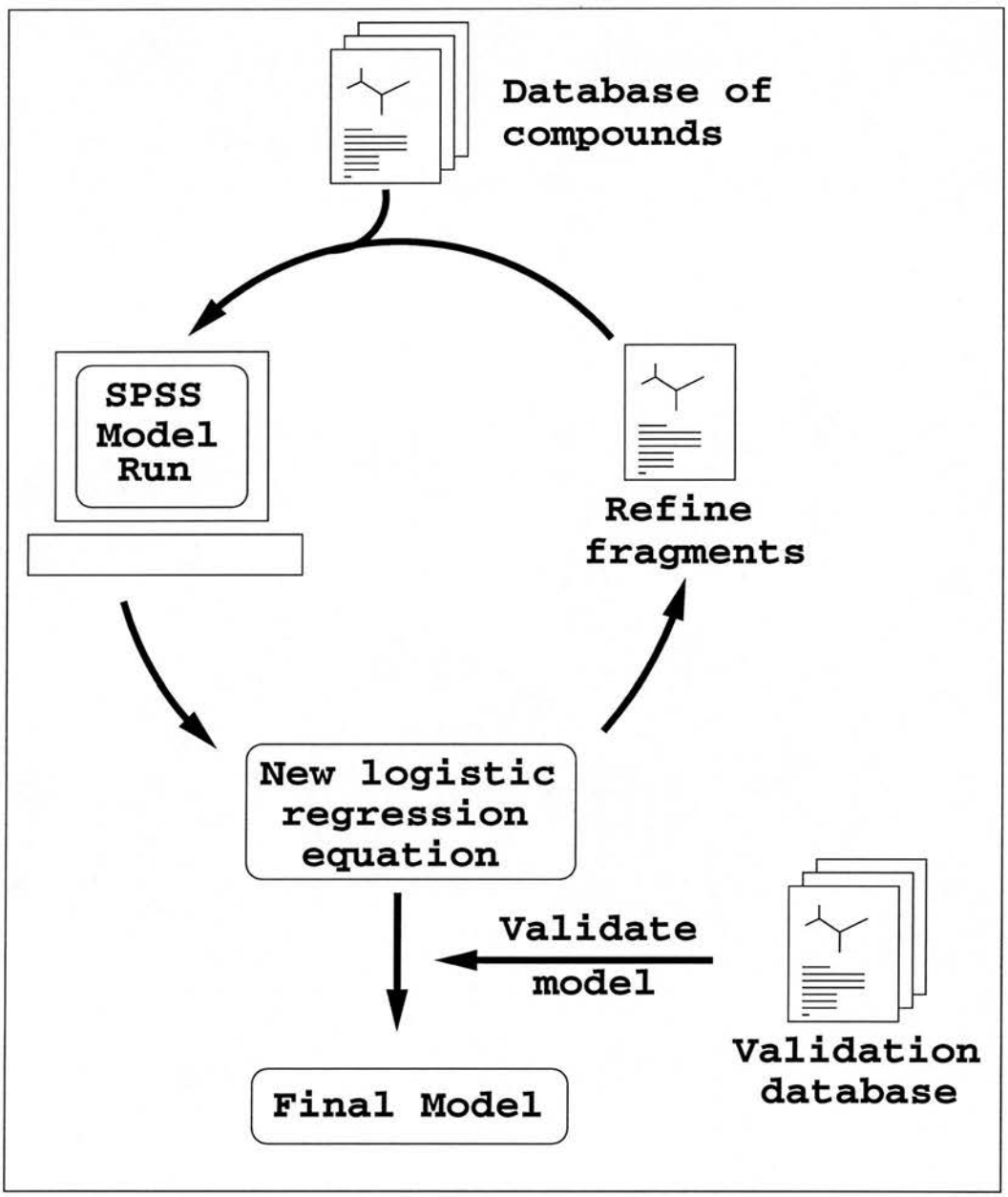


Figure 2.8: Logistic regression model was refined by adjustment of the fragment variables over a succession of model runs.

		Sensitizer status (S)		
		y	n	Total
Test	y	a	b	a+b
status (T)	n	c	d	c+d
Marginal Totals		a+c	b+d	

$$Sensitivity = Probability(T^{+}|S^{+}) = \frac{a}{(a+c)}$$

$$Specificity = Probability(T^{-}|S^{-}) = \frac{d}{(b+d)}$$

Table 2.10: Sensitivity and specificity. (After Hennekens and Buring [1, page 331]).

threshold to a lower or higher value one can alternatively increase sensitivity (and decrease specificity) or increase specificity (and decrease sensitivity). To quantify sensitivity and specificity, values were calculated using the formulae described in Table 2.10.

Creation of A Final Model

The final prediction program was developed from the existing fragment tabulating programs. The program takes a MDL molfile format connection table and calculates a likelihood of the compound being an asthmagen using a logistic regression equation.

The interface to the program was web based. The program is accessed via a web page into which the molfile of the chemical in question is pasted. The web based interface allows the program to be accessed from any computer architecture that supports a web browser (indeed a graphical user interface to the web browser is not essential). An additional advantage of a web based interface is that users do not have direct access to the program and that the program can be readily up-

dated as improvements are made. Finally, the possibility to monitor and restrict access to the program is available using this method.

The final model has been implemented as a world wide web (Internet) accessible program on a SGI $Indigo_2$ web server. The hazard prediction program was written in ANSI standard 'C' code and implemented as a cgi-bin program on the web server. Its URL is

```
http://www.bch.ed.ac.uk/cgi-bin/hazassess
```

and the W3 method used is the 'POST' method. The model is also available on the cdrom (see Appendix B).

The model may be used by users from any Web capable operating system however the creation of the molfile connection tables is dependent on a suitable chemical structure drawing package unless the molecules are small in which case a connection table could be manually created [1].

The user of the hazard assessment program is expected to create a molfile connection table for their molecule (See Figure 2.4). Full instructions of where to obtain the appropriate software are included with the package.

An interface was designed to allow standard molfile connection tables to be pasted into a HTML Form (see URL

```
http://www.bch.ed.ac.uk/~james/progs/
```

[2]

A text input was preferred because it allows access to the prediction program from non-graphical browsers. The model returns a hazard index for the compound and checks the database to determine whether the compound was an asthmagen or control in the original dataset.

[1] Users of Windows (3.11, 95, 98, NT) and Macintosh users can download copies of MDL's IsisDraw package which is available without charge to academic or individual home use from URL http://www.mdli.co.uk/

[2] Use the case-sensitive username **Guest** and the password **let_me_in** to access the programs.

Chapter 3

Results

Overview

The results will be presented in five parts: 'collected data'; 'clustering'; 'hazard odds ratios'; 'logistic regression'; and 'validation'. The data collection part will summarise the amount of data collected and present important epidemiological relationships noted. The clustering will be dealt with briefly as the methodology was dropped before the skin sensitizers were analysed. The hazard odds ratio and logistic regression section will each be subdivided into two subsections: the first and major subsection will look at how asthma compounds were differentiated from control compounds; and the second subsection will cover the results of methods used to differentiate between asthmagens and skin sensitizers. The final part, the validation, will include the results from the validation of the predictive model (developed from the logistic regression modelling to identify occupational asthma (OA) hazard).

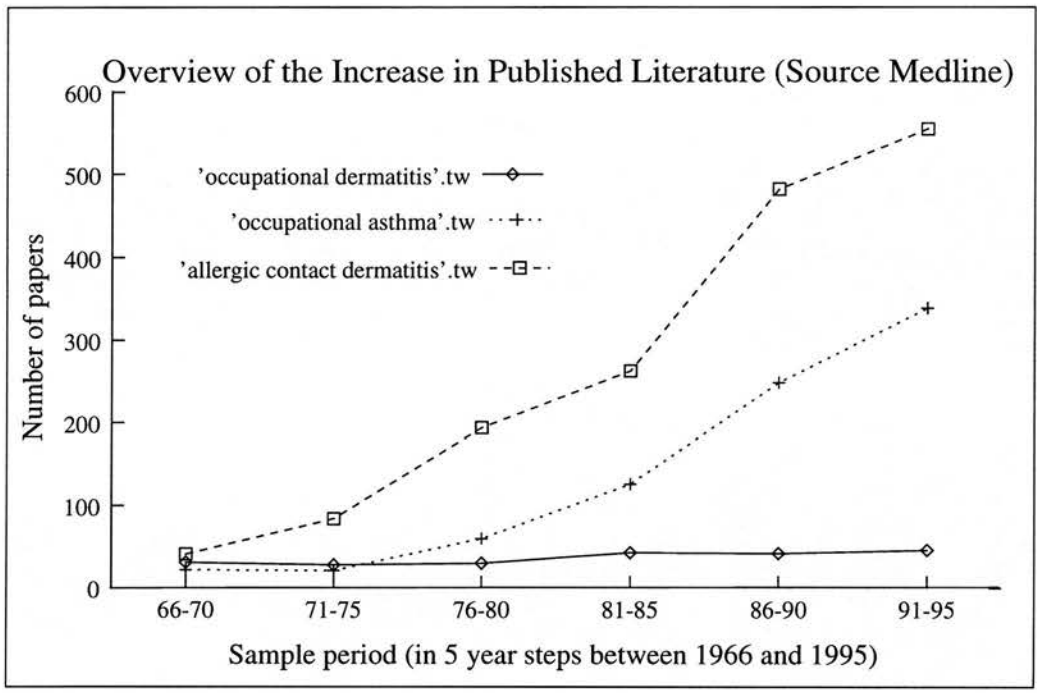


Figure 3.1: Number of papers yielded per 5 year period using particular textword search phrases in the MedLine database.

Category of Compound	Number of Compounds	Molecular Mass Median	Lower	Upper
Controls	302	118	89	187
Asthmagens	82	211	66	327
Skin sensitizers	200	207	143	290

Table 3.1: Number of chemicals and molecular mass distribution in each category of compounds.
(Upper and lower values refer to quartiles.)

3.1 Collected Data

The number of published papers using the phrases 'occupational asthma' and 'allergic contact dermatitis' has markedly increased over the last three decades (see Figure 3.1). Over the same period there is infrequent use and no notable increase of the phrase 'occupational dermatitis'.

In this period 1966-1994 a total of 82[1] asthmagens (see Appendix C.1) were identified by critical appraisal of published literature which satisfied the study criteria (see p50). CAS Registry Numbers were found for 77 of the active compounds and a structure was identified for all of them. Some compounds reported in the published literature as asthmagens failed to satisfy the full study criteria needed to be included in the analysis (see Appendix C.1). In addition 200 skin sensitizers[2] were identified and tabulated (see Appendix C.2). These data along with molecular mass distributions are summarised in Table 3.1.

[1]Since there was an overlap between the data collection and the preliminary data analyses some of the results that follow will indicate that less than 82 asthmagenic compounds were used. This occurred solely because all asthmagens had not at that point been identified and verified as satisfying the study criteria.

[2]A limit of 200 skin sensitizers was set for the purposes of the study although more than 200 occupational skin sensitizers are known.

Clinical and Epidemiological Data

Data deemed to be of potential epidemiological relevance was tabulated and database created. An example of the data available for each compound is given in Table 3.3. Several data fields have not been included as no data was found for this particular compound. These include symptom fields for cough, chest tightness, conjunctivitis, nasal symptoms, headache. For a large number of the fields data was too sparse to allow effective evaluation.

A large proportion of the asthmagens (31) were identified as hazards on the basis of single case reports, with over half the compounds (48) having 3 or fewer cases recorded. Twenty-two compounds had between 3 and 33 cases reported and eight compounds had in excess of 33. The commonest literature case reports of OA were due to toluene di-isocyanate, plicatic acid (from *Thuja plicata*, Canadian Western Red Cedar and *Thuja occidentalis*, Eastern White Cedar), phthalic anhydride, formaldehyde and abietic acid (found in colophony). These prevalence data represent minimum values.

Reported evidence for an IgE-mediated response was found in only 23 compounds, of which only 17 reported specific IgE (see Table 3.4). The median molecular mass for the 17 compounds for which specific IgE was noted was 206 (inter-quartile range 164-422). This was not significantly[1] different from the median molecular mass (265.91, 98-349) of the 11 compounds for which no evidence of an IgE mechanism was found when sought.

The type of asthmatic response noted varied greatly (see Table 3.5). Five compounds (azodicarbonamide; hexamethylene di-isocyanate; toluene di-isocyanate; plicatic acid; and a trimer of hexamethylene di-isocyanate) were reported as producing all three response types - immediate, late

[1]Significance tested using a non parametric test, the two sample Mann-Whitney test as described Eason *et al. [169]*

Name:	Tetrachloroisophthalonitrile	
CAS Registry No.:	1897-45-6	
Synonyms:	Chlorothalonil; 2,4,5,6-Tetrachloro1,3-benzene-dicarbononitrile; Daconil; Dac-2787; Bravo; Forturf; ExothermTermil; Termil	
General Notes:	USE - A fungicide, bactericide, nematocide. Agricultural and horticultural fungicide. [MERCK]	
Sensitisation notes:	Single case report by Honda et al.. Atopy defined by positive scratch test to just one of several common allergens. No elevated IgE was found. Using ELISA no specific IgE was found either.	
Category	Asthmagen?	Yes
	Skin sensitizer?	Yes
Physical Data	Physical form:	solid
	Melting point:	250 °C
	Boiling point:	350 °C
	Vapour pressure:	No data available
	Water solubility:	Practically insoluble
Epidemiological Data	Asthma absolute prevalence	Single case
	Asthma % prevalence:	Not available
	Basis of diagnosis:	Challenge tested
	Asthma response type:	Dual
	Sensitising concentration:	< 1μmole/litre of air
	IgE involvement:	No evidence found
	Latency:	1 year
	Persistence of asthma:	Not available
	% of atopic cases:	100%
Other symptoms:	Wheeze:	Yes
	Dyspnoea:	Yes
	....other symptoms	

Table 3.3: Example of a Database Entry.

Evidence for an IgE Mechanism	Compounds
IgE sought but not found	Furfuryl alcohol; Ethylene-diamine; Hydralazine; Isophorone di-isocyanate; Ethylene oxide; Tetrachloroisophthalonitrile; Abietic acid; Captafol; Hexachlorophene; Biuret of hexamethylene di-isocyanate; Dobutamine hydrochloride
Elevated levels of Total IgE	Hydroquinone; Carmine; Tetracycline; Tetrazene; Adipic acid; Methyl blue
Specific IgE found	Phthalic anhydride; Trimellitic anhydride; Himic anhydride; Hexahydrophthalic anhydride; Tetrachlorophthalic anhydride; Piperazine; Isoniazid; Hexamethylene di-isocyanate; Toluene di-isocyanate; Chloramine-T; 1,3,5-tris-(6-isocyanato-hexyl)-<1,3,5>triazinane-2,4,6-trione ; Plicatic acid; Black GR Reactive Dye(BK-5); Orange-GR Reactive Dye; Red-BBN Reactive Dye ; Phenylglycine acid chloride; MM22383;

Table 3.4: Asthma Compounds and Associated Immunoglobulin-E (IgE) Mechanism.
Note: Assignment of a compound to a category in the table indicates a minimum for information available. For example the observation that non-specific levels of IgE are elevated does not preclude there being specific IgE present. For many of the compounds the assignment is based on only one or two papers.

Types of Asthmatic Response Seen	Compounds
Immediate	3-Carene; Styrene; Trimellitic anhydride; Isoniazid; Amprolium hydrochloride; Chlorhexidine; trientine; 7-Amino cephalosporanic acid; Tetracycline; Hexachlorophene; Methyl blue; Cephalexin; Phenylglycine acid chloride;
Late	Formaldehyde; Furfuryl alcohol; Glutaraldehyde; Methylmethacrylate; Hydralazine; Isophorone diisocyanate; Methyl-2-cyanoacrylate; α-Methyl DOPA; Salbutamol; Acetic acid; 6-Amino penicillamic acid; Ampicillin; Benzyl penicillin; Cimetidine; Sodium iso-nonanoyloxybenzene sulphonate; Hexamethylene tetramene; Tetrazene; Glycyl compound; Fenthion
Immediate or Late	Paraphenylene diamine; diphenylmethane diisocyanate; Aminophylline
Dual	Himic anhydride; Hexahydrophthalic anhydride; Tetrachlorophthalic anhydride; Dimethylethanolamine; Tetrachloroisophthalonitrile; Captafol; Carmine; Pauli's Reagent; Sulfathiazole; Adipic acid;
Immediate or Dual	Pyromellitic di-anhydride; Abietic acid; Biuret of hexamethylene di-isocyanate; Black GR Reactive Dye (BK-5); Orange-GR Reactive Dye; Red-BBN Reactive Dye;
Late or Dual	Maleic anhydride; Phthalic anhydride; Ethylenediamine; Piperazine; Aminoethylethanolamine; 1,5-Naphthalene diisocyanate; Ethyl cyanoacrylate; Chloramine-T;
Immediate, Late or Dual	Azodicarbonamide; Hexamethylene di-isocyanate; Toluene diisocyanate; Plicatic acid; 1,3,5-tris-(6-isocyanato-hexyl)-<1,3,5>-triazinane-2,4,6-trione;

Table 3.5: Asthmagens by Asthma Response Type(s).
Note: Assignment of a compound to a category in the table indicates a minimum for information available. For many of the compounds the information is based on just one or two papers.

and dual responses albeit in different cases. A late response was observed (in at least some of the cases) with over half the compounds. A late response alone was seen with a quarter of the compounds. Only a third of the compounds caused immediate responses, and of these less than a fifth produced only an immediate response. Notably, no compounds for which specific IgE was observed caused solely a late asthmatic response. Compounds that did cause late responses for which specific IgE was noted also caused dual type responses[1].

Specific IgE was consistently found among the acid anhydride and reactive dye classes of asthmagens when sought. Furthermore, dual responses were noted (in at least some patients) for all acid anhydrides and reactive dyes.

The symptoms associated with sensitisation varied with chemical group: persons sensitized to acid anhydrides exhibited symptoms of wheeze, dyspnoea, rhinitis/rhinorrhea and cough. Skin and mucus membrane irritation were also noted though these are known irritant symptoms of acid anhydrides. Evidence for an IgE mechanism (either as raised serum levels or identification of specific IgE) was present in most cases of asthma due to anhydrides. Isocyanate asthma presented with symptoms of wheeze, dyspnoea, chest tightness, cough, rhinitis, runny eyes and skin irritation.

Nasal symptoms, described variously as rhinitis or rhinorrhoea were evident (in at least individual) for over half (38) the sensitising compounds. Other less common symptom, for example headache, were noted with formaldehyde, glutaraldehyde, triethylene tetramine, and hexamethylene tetramine.

[1]This is not to say immediate or late responses were not also seen.

3.2 Clustering

Prior to the start of this study some collaborative work between Dr Agius (The University of Edinburgh) and Dr Geoff Downs (formerly with The University of Sheffield, now with Barnard Chemical Information Ltd.) had been undertaken looking at clustering patterns of a dataset of asthma and control compounds. The first part of the clustering results relate to this collaboration. Clustering was performed by Dr Downs using software that was not readily available. The second part illustrates how clustering was used on the complete dataset. The clustering was performed using a C program written for the purpose[1]. The program and the output of the clustering can be found on the CD-ROM (see Appendix B). Clustering as a method was used only briefly in this work as it was super-ceded by other methods. It was not used to distinguish between skin and respiratory sensitizers.

The preliminary dataset contained 29 asthmagens and 29 controls (see Table 3.6) [164]. One control (tetrachloroisophthalonitrile) was later reclassified as an asthmagen. The results of clustering this dataset using different cluster sizes and methods are shown in Table 3.6. The clustering was performed by Dr G. Downs (Barnard Chemical Information Ltd.) using Ward, Guenoche, and Jarvis-Patrick clustering methods. The clustering was performed blind - Dr Downs did not know which compounds were asthmagens and which were controls. The Guenoche algorithm appeared to be effective with this dataset (particularly with few clusters) and, after consultation with Dr. Downs, the Guenoche method was adopted as the method to use for the remainder of the study.

[1]Dr Down's program was too expensive to purchase.

Controls	Asthmagens
Benzene	Δ3-Carene
n-Heptane	Styrene
2-Ethoxyethyl-acetate	Furfuryl alcohol
Butan-2-one	Formaldehyde
2-Methoxy ethanol	Glutaraldehyde
Acrylaldehyde	Maleic anhydride
Dibenzoyl peroxide	Himic anhydride
2-Hydroxypropyl acrylate	Phthalic anhydride
Dimethoxy methane	Trimellitic anhydride
Oxalic acid	Hydroquinone
Hexan-2-one	Methyl methacrylate
Butyl acetate	Pyromellitic di-anhydride
ortho-Toluidine	Ethylene diamine
Dimethyl amine	Paraphenylene diamine
Ethyl amine	Piperazine
Hydrazine*	Isoniazid
1-Methyl-2-pyrrolidone	Aminoethyl ethanolamine
1-Nitropropane	Hexahydrophthalic anhydride
Carbofuran	Azodicarbonamide
2,4,6-Trinitro toluene	Hexamethylene di-isocyanate
Nitro aniline	2,4-Toluene di-isocyanate
Nitro methane	α-Methyl DOPA
1,2-Dinitro propane	N-Methyl morpholine
ortho-Acetyl salicylic acid	Methyl cyanoacrylate
2,2'-Imino diethanol	Ethyl cyanoacrylate
Dimethyl formamide	Diethyl ethanolamine
2-Pyridyl amine	1,5-Naphthalene di-isocyanate
Diphenyl amine	Isophorone di-isocyanate
	2-Dimethyl aminoethanol
	Methyl-phenyl di-isocyanate

Table 3.6: List of control and asthma compounds used by Dr. G. Downs for the cluster analyses. .
Originally the data set consisted of 29 asthmagens and 29 controls but one of the controls was later found to be an asthmagen. * Note: Hydrazine does not contain carbon and so does not appear elsewhere in this thesis.

No. of Clusters	Ward	Guenoche	Jarvis-Patrick
10	0.19	0.50	0.31 (9)
15	0.57	0.56	0.47 (16)
25	0.70	0.77	-
30	0.70	-	0.70 (30)

Table 3.7: Kappa values for predicting asthma hazard.
The κ values derive from predictions made using the Ward, Guenoche and Jarvis-Patrick methods of clustering. The number of clusters cannot be selected using the Jarvis-Patrick method so the κ for nearest number of clusters is given (in brackets). The set of data included 31 asthmagens and 29 controls (one of the controls was found to be an asthmagen after the compounds were initially selected).

The final dataset for clustering consisted of 75 asthmagens and 302 controls. A Guenoche clustering of the 75 asthmagens of the complete set was performed using the descriptors of the type 2 fragment set (See p62). The clustering was performed twice: once with descriptors as continuous[1] variables; and then with descriptors coded in binary[2] form.

Figure 3.2 shows the results of clustering 14 times (to produce 15 clusters) using the continuous method. The values in parenthesis are the asthmagen:control ratio. It can be seen that the clustering is not even, indeed cluster 2 contains 207 of the original 377 compounds. The starting ratio of asthmagens to controls was approximately 1:4. A similar ratio would be expected in the clusters if the method was *not* distinguishing between asthmagens and controls. Several clusters, including cluster 2, appear to markedly differ from this ratio. Clusters 2 and 8 are 'control' clusters. The 'asthmagen' clusters such as 1, 3, 5, 7 and 9 have very few compounds. The kappa value for this clustering set was 0.38.

Performing the same cluster procedure as above on the same data but

[1]The fragment *frequency* of occurrence in each molecule (rather than just fragment presence or absence) was a determining factor in the clustering.

[2]A logical form where a 0 means the fragment is absent and a 1 indicates presence. This form is simplistic in that it only recognises presence of a fragment in a given structure and not the number of occurrences.

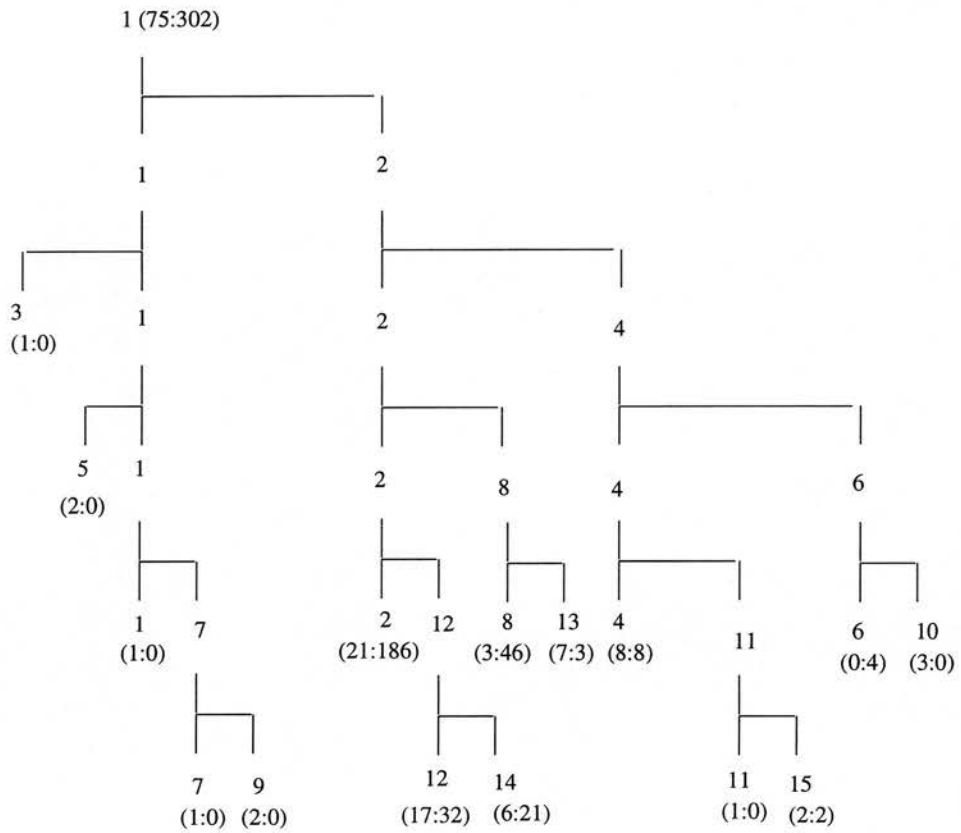


Figure 3.2: Hierarchical Structure of the Guenoche Continuous Cluster
The asthmagen:control ratio is given in parenthesis.

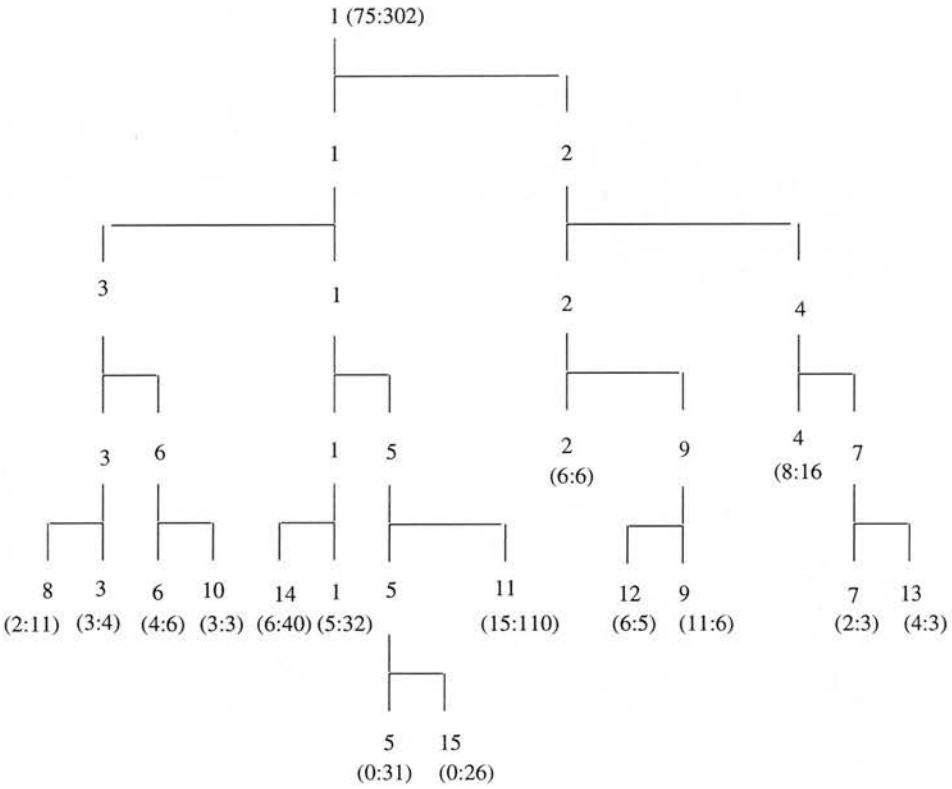


Figure 3.3: Hierarchical Structure of the Guenoche Binary Cluster
The asthmagen:control ratio is given in parenthesis.

using a binary cluster method a slightly more even spread of data (see Figure 3.3). The largest cluster (cluster 11) contains 125 compounds. Clusters 5 and 15 contain only control compounds. Cluster 5 has at least one halide in 29 of its 31 compounds. The only strong 'asthmagen' cluster is cluster 9. Although the binary cluster method produced a more even clustering, it produced fewer 'asthmagen' clusters than the continuous method. The kappa value for this clustering set was 0.41.

The compounds contained in each cluster are available on the CD-ROM (see Appendix B).

A number of cluster runs were performed using both the binary and the continuous method but varying the number of clusters. The results are shown in Figure 3.4. There is little difference between the two methods

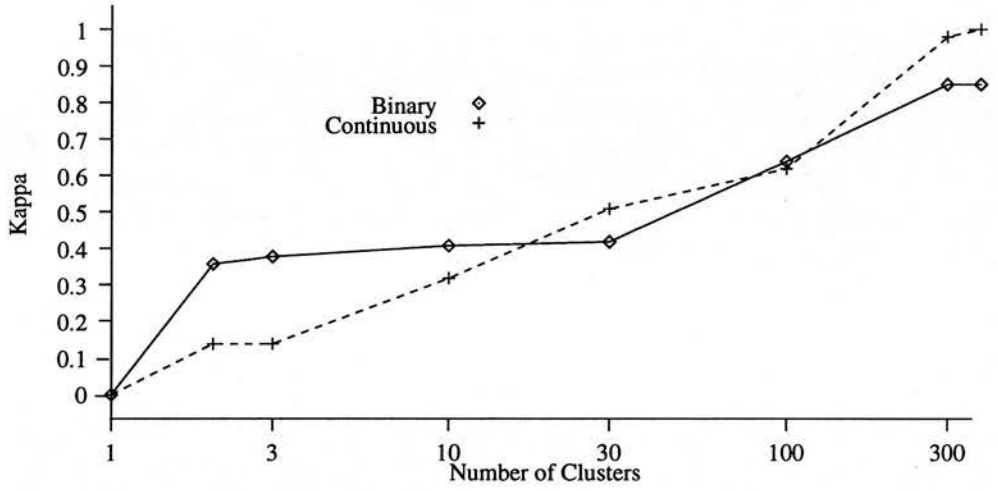


Figure 3.4: Effect of Cluster Size on the Kappa Value of Guenoche Clustering of 75 asthmagens and 302 controls.

except when fewer clusters are selected. When 15 or fewer clusters are selected the binary method produces the better kappa value.

3.3 Hazard Odds Ratios

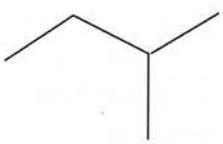
Chemical Substructure Fragments Associated with Asthma Hazard

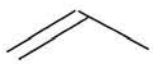
Seventy-six known respiratory sensitizers (asthmagens) were compared with 303 control compounds. Using the data contained in the fragmentation tables, Hazard Odds Ratios (HOR'S) were calculated (see page 73). The odds ratio for a particular frequency of occurrence n is given by the odds of a compound containing n fragments (of a specified type) being asthmagenic over the odds of a compound containing no fragments (of the same type) being asthmagenic.

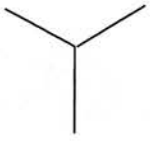
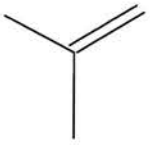
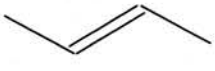
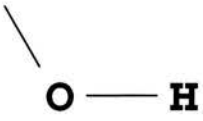
The HOR's and 95% C.I. for 130 Type 1 fragments (see p62) were calculated automatically. The results are shown in Table 3.8. The HOR's refer to discrete fragment occurrence frequencies.

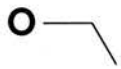
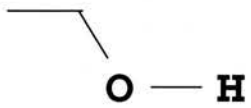
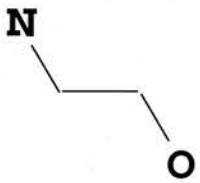
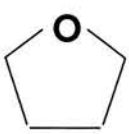
There is evidence that hazard increases with occurrence frequency in the compound for certain fragments (see fragments 7, 14, 15,16, 18, 27, 28 and 30 in Table 3.8). These fragments predominantly contain oxygen and / or nitrogen atoms. Several fragments have high HOR's for a single occurrence - these include the ethanolamine fragment (19), the 1,2-diaminoethane fragment (21), the furan fragment (20), the azine fragment (25), the thio-ether (37) and the acid anhydride fragment (34). The carbon-carbon single bond fragment (1) produces a statistically significant HOR at a frequency of 5 in a compound.

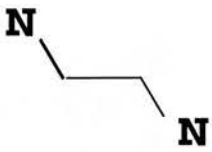
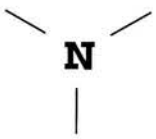
Table 3.8: Hazard Odds Ratios for Type 1 Fragments. An 'E' next to the 95intervals indicates an exact method was used to calculate the confidence intervals

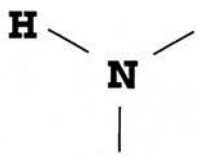
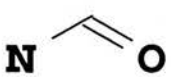
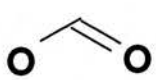
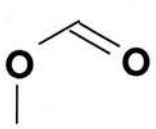
Atom or fragment	n	Odds Ratio	95% Confidence Interval Range
	5	10.37	3.19-43.28 E
1			
	5	7.73	2.86-20.90
2			
	5	6.29	2.25-17.60
3			
	1	3.58	1.19-10.71
	4	14.30	1.11-753.56 E
4	5	6.36	2.12-19.04
	4	8.22	2.98-22.66

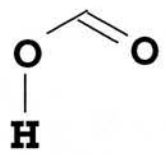
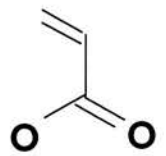
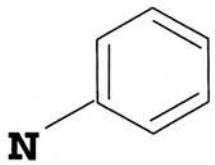
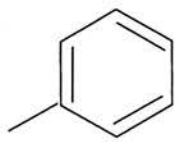
5		5	7.93	3.34-18.83
6		1	3.25	1.58-6.68
		2	3.36	1.39-8.12
		4	3.36	1.08-10.46
		5	9.24	3.46-24.70
7		1	2.79	1.20-6.52
		2	3.26	1.05-10.12
		3	2.68	1.40-5.14
		4	5.58	1.63-19.11
		5	5.06	2.10-1218
8		2	5.95	2.16-16.38
		5	3.42	1.91-6.12
9		1	3.42	1.08-10.83
		3	2.36	1.33-4.17
		5	3.85	1.61-9.19
10		2	2.33	1.08-5.02
		3	4.90	1.92-12.53
		5	9.52	3.93-23.07

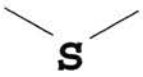
			
	2	6.57	2.19-19.71
11			
			
	1	5.37	2.60-11.08
12	2	4.07	2.03-8.16
	4	10.41	1.13-126.85 E
			
	1	5.31	2.22-1.71
13	3	2.27	1.12-4.62
	4	3.16	1.31-7.62
	5	6.81	3.03-15.31
			
	1	2.25	1.22-4.14
14	2	3.67	1.24-10.82
	3	16.50	1.27-869.57 E
	5	16.50	1.27-869.57 E
			
	2	3.39	1.15-9.96

15	3	15.25	1.18-803.68 E
			
	1	2.36	1.14-4.85
16	2	6.71	3.15-14.29
	3	14.53	4.40-47.98
	4	22.71	3.40-244.03 E
	5	72.67	8.81-3237.59 E
			
	3	3.54	1.18-10.64
17			
			
	1	1.94	1.02-3.66
18	2	3.42	1.16-10.04
	3	15.38	1.19-810.24
			
	1	17.41	6.18-49.07
19			
			
	1	12.90	2.22-132.03 E
20			

21		1	8.51	3.04-23.85
22	O=	2	7.33	3.51-15.32
23	S	1	4.27	1.93-9.47
24	=N	2	13.43	3.07-80.00 E
25	N=N	1	12.76	1.00-672.00
26		1	6.23	2.77-14.00

27		1 2	2.98 19.07	1.29-6.89 1.82-942.08 E
28		1 2	3.45 12.33	1.65-7.23 1.97-130.97
29		2	7.69	1.44-50.38 E
30		1 2	1.92 4.31	1.00-3.69 1.76-10.53
31		2	4.35	1.66-11.43

	1	5.55	2.37-12.96
32			
	1	3.36	1.13-10.03
33			
	1	26.26	3.07-1211.71 E
34			
	2	5.00	1.62-15.37
35			
	1	3.26	1.75-6.07
36			

				
		1	14.24	4.10-61.91 E
37				

Chemical Substructure Fragment	Environment	Hazard Odds Ratios	Upper & Lower 95% Confidence Intervals
Oxygen	Any	2.76	1.40-5.63
	Aliphatic	2.63	1.40-5.01
	Aromatic	2.11	1.04-4.76
Nitrogen	Any	5.56	3.12-9.95
	Aliphatic	6.29	3.53-11.24
	Aromatic	2.84	1.49-5.40
Sulphur	Any	4.26	2.05-8.82
	Aliphatic	3.68	1.66-8.09
	Aromatic	13.43	3.21-78.4
Chlorine	Any		
	Aliphatic	0.28	0.07–0.81
	Aromatic		

Table 3.9: Chemical elements with statistically significant hazard odds ratios.
(Asthmagens compared with controls).

The general HOR (for any occurrence) provides a useful measure of fragment hazard. Table 3.9 shows the HOR's for presence of atoms in either aromatic or aliphatic environment[1]. It should be noted that nitrogen atoms are more hazardous when in an aliphatic environment whereas sulphur atoms appear more hazardous in an aromatic environment. The hazard presented by oxygen atoms does not vary greatly between the two environments. The presence of halogens was associated with a lower than average OA hazard as halogen atoms were heavily represented in the control set. The HOR for the presence of aliphatic chloride atoms was 0.28 (0.07-0.81) and for aliphatic bromine was 0.00 (0.00-1.57, not listed in the table as it is not statistically significant).

When the HOR's for type 2 fragments are ranked in order, the fragments based on isocyanate, β-lactam ring, acid anhydride, ethanolamine

[1]An atom was deemed to be in an aromatic environment if it formed part of an aromatic ring or if one of its immediate neighbour atoms (to which it was bonded) was part of an aromatic ring.

Fragment	Hazard Odds Ratio	Upper & Lower 95% Confidence Limits
Isocyanate	∞	6.25-∞
β-lactam ring	∞	6.25-∞
Acid anhydride	30.64	3.71-1385
Ethanolamine backbone N-C-C-O	19.93	6.07-83.8
Acrylate derivative	13.04	2.25-135
Imine C-NH-C	8.24	3.47-19.8
Carboxylic acid -COOH	5.48	2.25-13.37
Any double bond X=X	3.77	2.01-7.17
Carbonyl	3.15	1.82-5.47

Table 3.10: Statistically significant hazard odds ratios for chemical substructure fragments.
These are Type 2 (see p62) fragments. (Asthmagens compared with controls)

and acrylate groups all present HOR's indicating a greater than tenfold increase in the likelihood that compounds containing them are an OA hazard (see Table 3.10).

HOR's for type 2 fragments were examined for increasing fragment occurrence rates. Significant HOR's were noted for both aliphatic and aromatic occurrences of the following hetero-atom types: nitrogen, oxygen, sulphur (see Table 3.11). The presence of two or more imine or carbonyl groups presented a significant OA hazard whilst a single occurrence was not significant .

Fragment	Fragment Occurrence Rate			
	1	2	3	≥ 4
Aliphatic oxygen	1.13	**2.33**	**5.41**	2.60
Any amine	**2.55**	**16.04**	**23.39**	∞
Aromatic carbonyl	2.46	**4.38**	∞	∞
Aromatic nitrogen	1.49	**3.27**	**9.81**	0.00
Ethene derivative	**2.79**	3.26	**2.68**	**5.58**
Imine derivative	2.52	**13.43**	∞	∞
Carbonyl	1.16	**7.33**	∞	∞
Any nitrogen	**2.35**	**6.71**	**14.53**	**22.71**
Any oxygen	1.18	**2.33**	**4.90**	2.03
Any double bond	1.29	**3.10**	**10.36**	**14.25**

Table 3.11: Hazard odds ratio by frequency of occurrence of Type 2 substructure fragments.
(Asthmagens compared to controls, statistically significant (at 5% level) HOR's appear in **bold** text).

Fragment	Hazard Odds Ratio	Lower & Upper 95% Confidence Intervals
Oxygen	2.47	1.06–6.44
Aliphatic Oxygen	2.42	1.13–5.51
Aromatic Sulphur	3.12	1.05–9.04
Carboxylic Acid	7.14	2.45–22.21
Aliphatic Carboxylic Acid	7.57	2.44–25.67
β-lactam	6.81	1.38–43.06
Acid Anhydride	24.57	3.00–1113.17

Table 3.12: Hazard odds ratios to differentiate between skin and respiratory sensitizers.

Using Hazard Odds Ratios to Differentiate between Skin and Respiratory Sensitizers

Fragment tables were created for the 58 asthmagens (which were not skin sensitizers) and 179 skin sensitizers (which were not asthmagens). Hazard odds ratios and χ^2 tests for trend were calculated for the explicit set of 48 Type 2 fragments plus an additional X-C-C-X fragment (where X could be oxygen or nitrogen[1]).

Seven fragments had significant HOR's for occurrence (see Table 3.12) and thirteen fragments had statistically significant linear trends (p value $\leq$0.05, see Table 3.13). A strong linear trend is seen with the carboxylic acid fragment. All of the listed HOR's had a significant p value for trend except the *aromatic sulphur* fragment . It should be noted that the presence of an acid anhydride fragment is nearly 25 times more likely to occur in a asthmagen than a skin sensitizer (Table 3.12).

[1]This is an example of a fragment that could not be represented by the type 1 method.

Fragment	Trend	Significance
Aliphatic Nitrogen	4.05	$p = 0.044$
Oxygen	8.12	$p = 0.004$
Aromatic Oxygen	3.97	$p = 0.046$
Aliphatic Oxygen	6.05	$p = 0.014$
Aromatic Chlorine	6.96	$p = 0.0084$
Carboxylic acid	20.16	$p = 0.00001$
Aliphatic carboxylic acid	19.43	$p = 0.00001$
Any double bond	6.41	$p = 0.011$
C=N (double bond)	6.54	$p = 0.011$
C=O (double bond)	5.81	$p = 0.016$
Aromatic C=O	4.51	$p = 0.034$
β-lactam	9.04	$p = 0.0026$
Acid anhydride	17.10	$p = 0.00004$

Table 3.13: Linear trend for fragments in the model differentiating between the skin and respiratory sensitizers.
A positive trend for a fragment favours classification as an asthmagen as the occurrence of that fragment increases.

3.4 Logistic Regression

Logistic regression models were created from the fragment data using a backward stepwise method with each iteration being tested by the *log likelihood ratio* statistic (see p77). Three models will be described here: the first model is based on the asthmagen and control dataset and it was used for the majority of the plotted data; the second model was used to predict between skin and respiratory sensitisation; finally, the third model is used in the programs available on the accompanying CD-ROM (see AppendixB). Results will refer to the first model unless stated otherwise.

The first logistic regression model was created using a dataset of 75 asthmagens and 302 controls. The logistic regression program had a limit of 50 on the number of variables that could be entered for logistic regression analysis. This precluded the use of the type 1 fragment dictionary and so selected type 2 fragments were used. The predictor fragments and their coefficients are shown in Table 3.14. Positive β coefficients indicate that a fragment is associated with increased asthma hazard. Negative β coefficients indicate that a fragment plays a 'protective' role. The standard error is a measure of how wrong the fragment coefficient can be. A large β coefficient will generally produce a large standard error unless a fragment has decisive effect on activity. The model has two large magnitude β coefficients (for the ketone and the aldehyde fragment groups). The standard error for these two coefficients is also very large. The acid anhydride and the isocyanate fragments also have large β coefficients but the standard errors are relatively small. The significance of the log likelihood ratio is the parameter which determines whether a fragment remains in the model[1]. Note the aldehyde fragment has the least significant log likelihood ratio.

[1]All fragments start in the model with the backward stepwise method, they are then removed if their contribution is not significant.

Fragment	β coefficient	Standard Error	Significance of the log likelihood ratio
Carbonyl	14.36	164	0.0043
Isocyanate	5.74	21.6	< 0.0001
Acid anhydride	5.23	1.20	< 0.0001
Acrylate	2.81	0.99	0.004
Aliphatic carboxyl	2.00	0.58	0.0009
Amine	0.99	0.36	0.0037
Carbon	- 0.61	0.16	< 0.0001
Aromatic nitrogen	- 0.64	0.34	0.0410
Ether	- 1.00	0.46	0.0148
Benzyl	- 1.02	0.48	0.0305
Aliphatic oxygen	- 1.11	0.28	< 0.0001
Aromatic chlorine	-2.07	0.49	< 0.0001
Aliphatic chlorine	- 2.27	0.51	< 0.0001
Aliphatic sulphur	- 2.36	0.66	0.0004
Phosphorus	-3.24	1.28	0.0009
Amide	- 3.46	1.45	0.0053
Aliphatic fluorine	-7.13	23.3	0.0035
Aliphatic bromine	-9.93	24.7	< 0.0001
Aldehyde	-12.21	164	0.0561
MOLECULAR MASS	0.066	0.01	< 0.0001
CONSTANT α	-4.30	0.5	–

Table 3.14: Logistic regression model differentiating between asthmagens and control chemicals.
The variables α and β correspond to the constant and coefficient of the logistic regression equation. See Appendix D.2 for details of the values shown.

The hazard prediction given by the logistic regression model for a particular compound is dependent on the correct classification of the compounds in the learning dataset. The reliability of this classification may be related to the 'method of diagnosis' and/or the 'prevalence' of OA to each learning dataset compound. Hazard predictions based on the model were made for each compound in the learning dataset and these predictions were then plotted against 'method of diagnosis' (see Figure 3.5) and 'prevalence' (see Figure 3.6). Figure 3.5 demonstrates that the model prediction is higher for those compounds identified in the literature as asthmagenic on the basis of respiratory challenge. Similarly, Figure 3.6 demonstrates that model predicts lower hazard values for those compounds cited as asthmagenic in single case reports only.

A number of compounds were misclassified by this first regression model. The active compounds that were most seriously misclassified (that is with predicted hazard less than 0.25) were: 2-diethyl-ethanolamine (predicted hazard 0.22); styrene (0.02); para-phenylene diamine (0.17); dimethylethanolamine (0.21); N-methyl morpholine (0.08); hydroquinone (0.04); 3-carene (0.09); 2-methyl-3,5-dinitro benzamide (0.18); captafol (0.20); phenyl glycine acid chloride (0.15); fenthion (0.15); acetic acid (0.12); ethylene oxide (0.01); chloroxylenol (0.08); and furfuryl alcohol (0.01).

The control compounds that were seriously misclassified (that is with predicted hazard greater than 0.75) were acetic anhydride (predicted hazard: 0.83); 2,2'-iminodi(ethylamine) (0.90); hexahydro-1,3,5-trinitro-1,3,5-triazine (0.83); methacrylic acid (0.77) and 6,6-di-*tert*-butyl-4,4-thiodi-*m*-cresol (0.90).

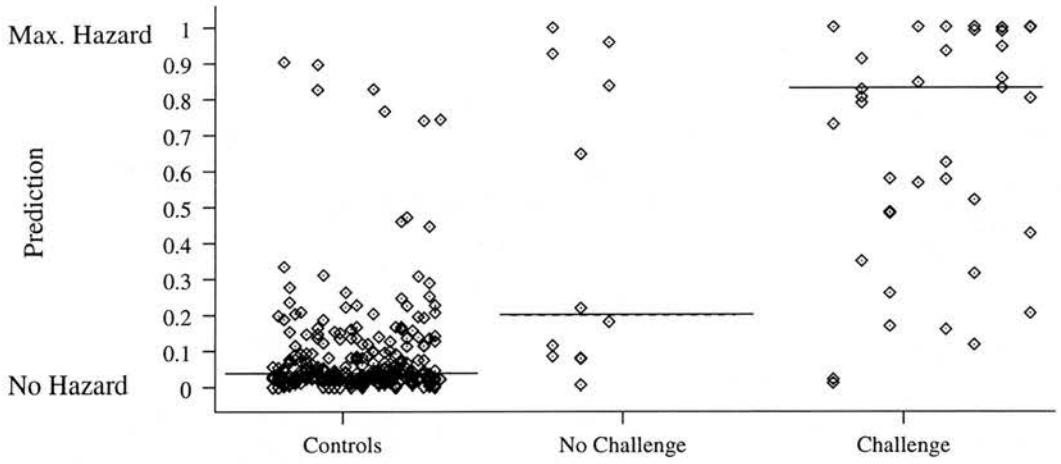


Figure 3.5: Scatter plot showing the predictive performance of the logistic regression model (by method of diagnosis of occupational asthma).
Each point represents one chemical. Horizontal lines indicate median value. The Y axis represents predicted hazard in the range 0 (no hazard) to 1 (maximum hazard). Challenge refers to whether diagnosis of a chemical as an asthmagen was based on a respiratory challenge test.

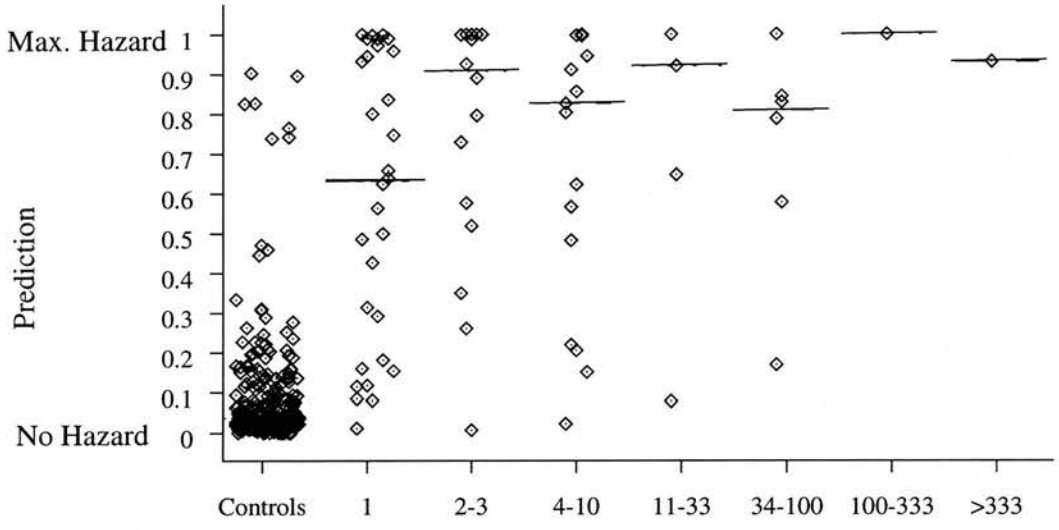


Figure 3.6: Scatter plot showing the predictive performance of the logistic regression model (by prevalence of occupational asthma).
Each point represents one chemical. Horizontal dotted lines indicate median value. The Y axis represents predicted hazard in the range 0 (no hazard) to 1 (maximum hazard).

		Predicted		
		Skin	Asthma	% Correct
Observed	Skin	169	11	93.89%
	Asthma	20	37	64.91%
		Overall 86.92%		

Table 3.15: Classification table for the logistic regression model differentiating between skin and respiratory sensitizers.

Distinguishing Between Skin And Respiratory Sensitizers Using Logistic Regression

In the same way that a logistic regression model was developed to distinguish between actives from controls for OA, it is possible to develop a model to discriminate between skin and respiratory sensitizers. This was done because it proved impossible to produce a valid set of controls for skin sensitizers. (There is no suitable equivalent of the airborne occupational exposure limits (OEL's, as seen in the EH40 documentation [161].)

Of the 75 asthmagens and 200 skin sensitizers used in this analysis there were 18 compounds which appeared in both sets. In addition there were a further 20 asthmagens for which skin sensitisation had been alleged by the asthma literature but the compound had not been cited in the skin sensitizer literature search. A logistic regression model was created to distinguish between the asthmagens and skin sensitizers in this set of 237 compounds. The discrimination of asthmagens using this model is shown in Table 3.15. The logistic regression model correctly identified 87% of the skin and respiratory sensitizers.

Table 3.16 displays the substructure fragments used by the model. The model predicts 1 for a perfect asthmagen and 0 for a perfect skin sensitizer. Consequently positive β coefficients favour an asthmagen classification. Those fragments (such as the acid anhydride, the aromatic nitro and any nitro) for which the magnitude of β (positive or negative) was

Fragment	β coefficient	Standard Error	Significance of the log likelihood ratio
Aliphatic carboxyl	3.00	0.65	<0.0001
Aliphatic chlorine	-0.65	0.43	0.0600
Aliphatic nitrogen	0.66	0.16	<0.0001
Aliphatic sulphur	-0.63	0.41	0.0725
Acid anhydride	10.00	18.48	<0.0001
Aromatic amine	-1.32	0.59	0.0091
Aromatic nitro	7.88	26.11	0.0449
Aromatic oxygen	0.72	0.28	0.0082
Aromatic sulphur	2.37	0.52	<0.0001
Isocyanate	0.95	0.59	0.0857
Nitro	-7.43	26.10	0.0641
X-C-C-X (X is O or N)	-0.25	0.11	0.080
Constant	-2.37	0.38	–

Table 3.16: Logistic regression model to differentiate between skin and respiratory sensitizers. and
Where β is the coefficient for the fragment, S.E. is the standard error. See Appendix D.2 for details of the values shown.

large were infrequently occurring fragments (with a large standard error (S.E) value). Figure 3.7 shows that the model does discriminate moderately well. The kappa value for the model is 0.62. Amongst the skin sensitizers which were predicted as asthmagens were aminotriazole (predicted 0.57), dicyanodiamide (0.57), acrylic acid (0.65), benzo-<d>-isothiazol-3-one (0.66), N,N-ethyl-4-toluene sulfonamide (0.66), Methylene bis(4-cyclohexylisocyanate (0.70 and oxacillin (0.77).

When used to predict sensitizer status of those compounds which are believed to cause both skin and respiratory sensitisation, the equation arising from the data in Table 3.16 produces some interesting results. The predictions range from 0.0067 (paraphenylene diamine) to 0.9995 (phthalic anhydride). The median value is 0.175 (ethylene-diamine and piperazine). They are plotted in the centre column of Figure 3.7.

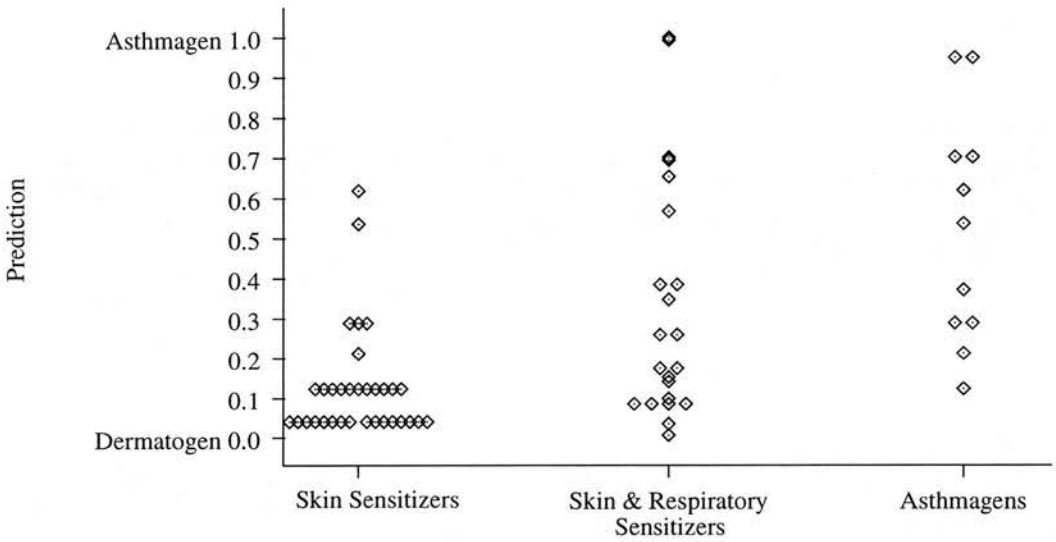


Figure 3.7: Distinguishing Between Skin and Respiratory Sensitizers Using Logistic Regression: Observed groups and predicted probabilities.
Each point represents 5 case chemicals. The Y axis represents predicted hazard in the range 0 (skin sensitizer) to 1 (asthmagen).

The Web-Based Model

The final logistic regression model was included in an Internet based utility to allow access from any suitable web browser. The user has to obtain a *.mol file to use the program but a chemical drawing package called IsisDraw is available for non-commercial use from the URL `http://www.mdli.co.uk/`. Access to predictions is via a web form (see Figure 3.8). The user 'pastes' in the text of the *.mol file into the form text entry area and then clicks on the 'Submit' button. The program returns a dynamically created web page incorporating the prediction for that compound (see Figure 3.9).

The final model was derived from 77 asthmagens and 302 controls. It produced a kappa value of 0.78. The equation of the model is:

$$y = \frac{e^{(-4.20+W)}}{1+e^{(-4.20+W)}} \text{ where } W = \begin{array}{l} -0.84(Carbons) \\ -1.68(Alkyl.Oxygens) \\ -3.86(Phosphorus) \\ -2.66(Alkyl.Sulphur) \\ -7.59(Alkyl.Fluorine) \\ -2.60(Aromatic.Chlorine) \\ -2.97(Alkyl.Chlorine) \\ -10.59(Alkyl.Bromine) \\ +4.49(Isocyanate) \\ +1.46(Alkyl.Carboxyl) \\ -1.18(Ether) \\ -12.06(Aldehyde) \\ +13.51(Ketone) \\ -0.82(Benzene) \\ +1.14(C.Double.Bond.O) \\ +2.41(Acrylate) \\ +4.09(Anhydride) \\ -4.63(Amide) \\ +3.20(Ethanolamine) \\ +0.08(Molecular.Weight) \end{array}$$

The fragment names in brackets should be substituted with the occurrence frequency of that fragment in the molecule under investigation. It should be noted that the definition of a fragment is embedded into the programs and the names assigned here may be used as an approximate guide.

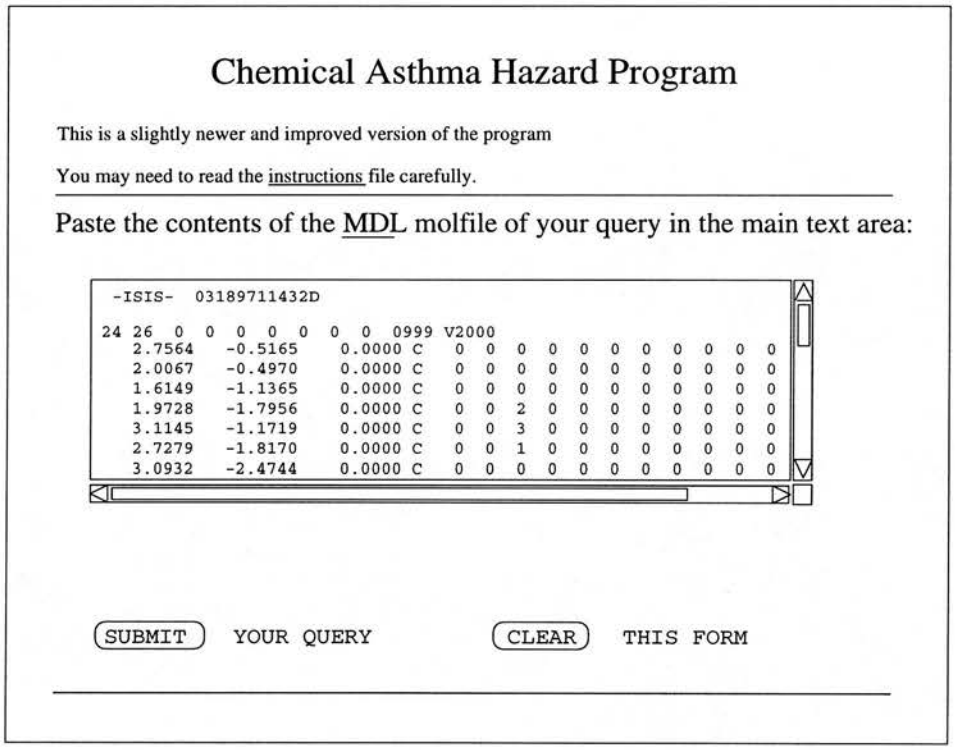
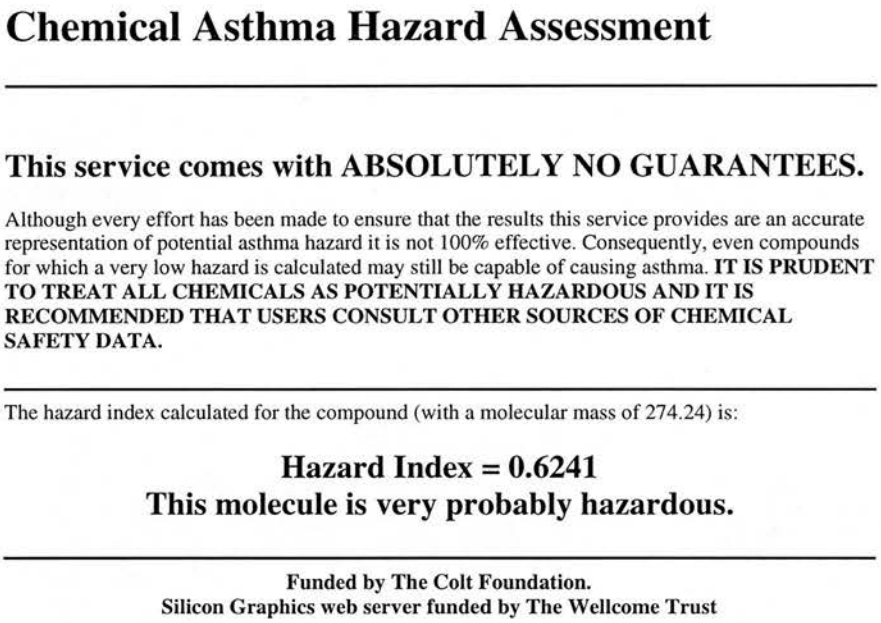


Figure 3.8: The Entry Form of Web-Based Asthma Hazard Prediction Program.



Chemical Asthma Hazard Assessment

This service comes with ABSOLUTELY NO GUARANTEES.

Although every effort has been made to ensure that the results this service provides are an accurate representation of potential asthma hazard it is not 100% effective. Consequently, even compounds for which a very low hazard is calculated may still be capable of causing asthma. **IT IS PRUDENT TO TREAT ALL CHEMICALS AS POTENTIALLY HAZARDOUS AND IT IS RECOMMENDED THAT USERS CONSULT OTHER SOURCES OF CHEMICAL SAFETY DATA.**

The hazard index calculated for the compound (with a molecular mass of 274.24) is:

Hazard Index = 0.6241
This molecule is very probably hazardous.

Funded by The Colt Foundation.
Silicon Graphics web server funded by The Wellcome Trust

Figure 3.9: The Result Returned from the Web-Based Prediction Program.

3.5 Validation

The prediction models previously described have been subject to internal validation. The study included an external validation with active and control chemicals not used (or known about) during the creation of the predictive model. The use of these data to validate the models is discussed in this section.

Only eight novel asthmagens (which satisfied study criteria) were identified from peer-reviewed literature post 1994 (See Table 3.17). So as to avoid observer bias the external validation set was generated by a different observer (Dr. R. Agius). The validation dataset (Dr Agius) included 62 SWORD[1] reported compounds. After rejection of repeated compounds[2], those that were not LMW organic compounds and compounds already included as asthmagens (See Table D.1), the data set contained 38 compounds (See Table 3.18). Fourteen of the compounds included were controls in the original dataset but were still included.

Of the twenty-five compounds originally identified as controls, nine were rejected because they were controls in the original dataset (See Table D.1).

[1]Surveillance of Work-related and Occupational Respiratory Disease, a scheme introduced in 1992 to allow chest physicians to report occupational groups and agents with a high risk of respiratory disease (including asthma).

[2]Several compounds were entered in the SWORD list under synonyms.

Compound	Hazard Index
Ceftazidime	1.00
Piperacillin sodium	1.00
1,3-bis(isocyanato-methyl)-cyclohexane prepolymer BIC	1.00
Indigotine	0.81
Triglycidyl isocyanurate	0.62
Penicillin G	0.47
Ethyl methacrylate	0.39
Triethanolamine	0.20

Table 3.17: Validation asthmagen compounds - identified from peer reviewed literature.

Compound Name	Hazard Index	Compound Name	Hazard Index
Cefotaxime	1.00	Halothane	0.00
Tetraethyl pentamine	1.00	Fenitothion	0.01
Cefuroxime	0.98	Dimethyl sulphate	0.01
Citric acid	0.92	1-Chloro-2,3-epoxypropane	0.01
Acetic anhydride	0.83	Toluene	0.02
Methacrylic acid	0.77	Diphenylmethane	0.02
Oleic acid	0.71	Dichloromethane	0.02
Pyridostigmine	0.66	Isophthalic acid	0.02
Diethylene diamine	0.58	Xylene	0.02
Codeine	0.55	Trichloroethane	0.03
Butyl methacrylate	0.49	Methylethyl ketone	0.03
Isobutyl methacrylate	0.49	Phenyl-β-naphthalene	0.04
Hydroxypropyl methacrylate	0.43	Cyclohexanone	0.04
Dolobid	0.42	Cyanogen chloride	0.04
Ketamine	0.34	Amyl acetate	0.05
Triethylamine	0.23	Butyl acrylate	0.05
Acetaldehyde	0.14	Lindane	0.07
Dimethylformamide	0.14	Butyl phthalate	0.08
Formic acid	0.12	Acetyl salicylic acid	0.08

Table 3.18: Validation asthmagen compounds - identified from the SWORD scheme.

The model was assessed quantitatively by calculation of a kappa value at different threshold values. The kappa value was derived from the hazard predictions given by the model for the compounds in the validation dataset. A clear demarcation between the controls and the peer reviewed asthmagens was noted with a maximal kappa value of 0.81 (threshold taken at 0.38, see Table 3.20 and Figure 3.10). The kappa value for distinguishing between controls and asthmagens identified by peer reviewed literature was approximately 0.7 over a range of threshold values (from 0.2 to 0.6). The SWORD data has been split to indicate the compounds reported in SWORD which were controls in the original (learning) dataset. Higher hazard prediction values are noted for those SWORD compounds which were not controls in the original dataset (median 0.34) when compared to those SWORD compounds which were

Compound	Hazard Index
Propanolol	0.43
Dinonyl phthalate	0.38
4,4'-methylenedianiline	0.22
Endrin	0.07
Aniline	0.05
Divinyl benzene	0.03
Diethyl sulphate	0.02
Ortho-toluidine	0.02
1-chloro-4-nitro benzene	0.02
1,2 Dichloro ethane	0.02
Propylene oxide	0.01
Paracetamol	0.01
Dimethyl ether	0.01
Isoflurane	0.00
Enflurane	0.00
1,1,1,2-Tetrafluoro ethane	0.00

Table 3.19: Validation control compounds - identified from known hazardous chemicals.
Although enflurane was entered in as a validation control compound a literature check on this compound revealed it had in fact been reported to cause occupational asthma[131]. Indeed, the report predated the exclusion date of 1/1/1995 so it should have been in the original set of asthmagens.

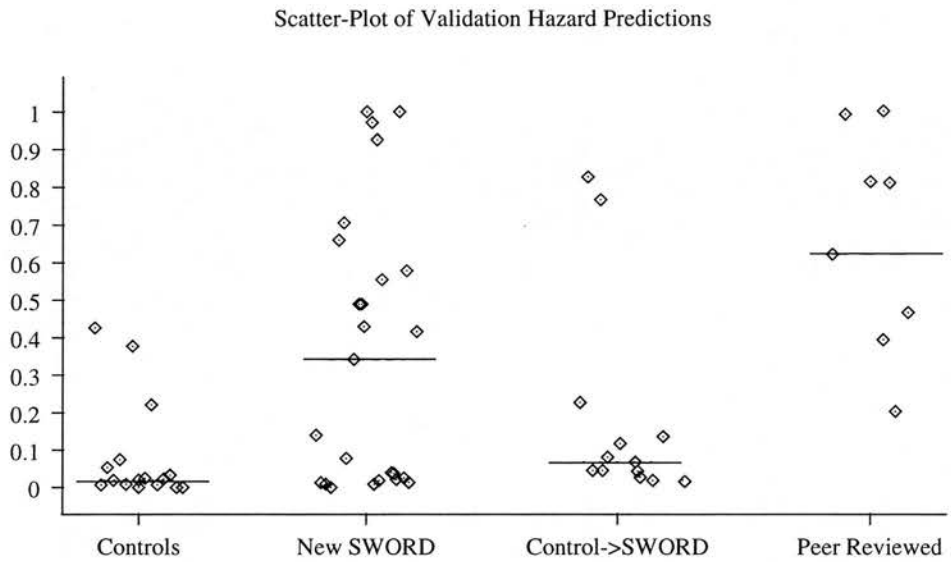


Figure 3.10: Scatter plot showing the predictive performance of the logistic regression model for the validation data.
Each point represents one chemical. The horizontal lines indicate median value. The Y axis represents predicted hazard in the range 0 (no hazard) to 1 (maximum hazard). The median values were 0.02 for the controls, 0.10 for the all SWORD cases and 0.71 (0.43-1.00) for the peer reviewed cases. 'New SWORD' includes only those SWORD compounds which *were not controls* in the original dataset. 'Control-> SWORD' includes only those SWORD compounds which were controls in the original dataset.

controls in the original dataset (median) 0.07). The median values for the asthmagens derived from peer reviewed literature and from controls were 0.62 and 0.02 respectively.

Threshold	0.2	0.3	0.38	0.4	0.5	0.6
All SWORD	0.17	0.20	0.22	0.22	0.17	0.14
New SWORD	0.30	0.36	0.32	0.37	0.34	0.20
SWORD & peer reviewed combined	0.24	0.24	0.26	0.25	0.20	0.18
Peer reviewed only	0.74	0.73	0.81	0.71	0.69	0.69
New SWORD & peer reviewed combined	0.39	0.41	0.42	0.39	0.30	0.25

Table 3.20: Kappa (κ) values for different threshold levels. 'New SWORD' indicates only those SWORD compounds which *were not controls* in the original dataset. 'All SWORD' includes those SWORD compounds which were controls in the original dataset.

Chapter 4

Discussion

Overview

The three aims of this thesis were firstly, the collection of data; secondly, the use of that data to describe (and predict) occupational chemical sensitizing hazard using chemical structure; and thirdly the use of that relationship to stimulate novel mechanistic hypotheses. The effectiveness of this work in addressing the first two aims will be considered first. The discussion will conclude by inferring from the observed results mechanistic hypotheses for the pathogenesis of occupational sensitisation to LMW organic chemicals.

4.1 The Collection of Data

The clinical, physiological and immunological investigation of cases of occupational asthma (OA), coupled with occupational hygiene assessments and epidemiology provide useful information about industrial chemical causes, and the relationship between exposure and response. However they do not permit easy general conclusions or predictions

about the risk that may be associated with novel chemical entities nor do they permit a characterisation of the diversity of chemical mechanisms that lead to asthma.

The structure-activity analyses described in this work were based on data obtained from a systematic literature search for compounds described as causing occupational skin or respiratory sensitisation. These data were obtained from published medical literature found in journals up to and including 1994. This section discusses the methods of collection and the reliability of the data.

The literature searching had to be extensive yet still follow strict criteria to minimise any selection bias. The choice of keywords (see Table 2.1) was therefore very important. The use of terms 'manufacturing' and 'industry' could be possibly viewed as introducing a bias towards non-service occupations. However, the rationale for using such keywords was to identify as many valid occupational asthmagens as possible. It could be argued that the control compound selection showed a similar bias in that compounds known to be hazardous to health are used predominantly in manufacturing industries. There are exceptions - workplaces such as hospitals would perhaps be under-represented as a result of the keyword. Any use of occupation or workplace-specific keywords to compensate for this bias could produce a greater bias. A consequence of this was that relatively non-specific keywords and an exhaustive scrutiny of several thousand titles and abstracts online was necessary to identify case compounds.

The effectiveness of the literature searching for the asthmagens is difficult to assess as there is not a definitive journal from which sufficient references can be exhaustively checked to ensure that search techniques work well. This was important as there are only a fairly limited number of suitable asthmagens for the study whereas it was easier to find sufficient numbers of skin sensitizers. For contact sensitizers, however the journal 'Contact Dermatitis' proved suitable for such an as-

sessment. In Volume 26 'Contact Dermatitis' 28 references were identified which identified cases of skin sensitisation by compounds which would be suitable for inclusion in the study. Of these 28 references, only 6 were identified by the keywords used in the literature search. It would have been necessary to enter occupation specific keywords to increase the sensitivity of the literature but this would have introduced selection bias (towards those occupations). Ultimately the choice of keywords is a balance between finding suitable numbers of compounds given that these need to be 'manually' scrutinised, and minimising selection bias. It is therefore probable many suitable case chemicals (for both skin and respiratory sensitizers) were not identified in this study.

One further problem associated with the use of keywords was spelling. The spelling of *sensitisation / sensitization* is just one such example. The work-around was to use both in textword searches. However this illustrates the need for an iterative search procedure.

The quality of the collected data, particularly for respiratory sensitisation, was extremely variable. The authors of early papers in particular did not have the facilities to do a full investigation of chemically-induced asthma cases. There were very few papers published covering OA case diagnosis techniques (see Figure 3.1) and so diagnostic methods were primitive. Furthermore the lack of computer-based journal indexes made searching for papers on similar topics far more difficult than it is today.

The diagnosis of causative agents by challenge testing (such as the tray tipping method for dusts and powders used by Pepys *et al.* [98]) helped improve the confidence with which chemicals could be described a cause of asthma. However many papers (including recent ones, e.g. Gadon *et al.* [93]) fail to evaluate OA cases fully and thus leave some doubt as to whether the true causative agent was discovered. The problem arises because physicians do not always get the patient's agreement to perform challenge tests, nor do they have access to sufficient

resources to do a full clinical work-up of each case. Therefore the confidence with which a chemical can be described as a sensitizer is based both on the number of cases and the methods of diagnosis. This culminates in the problem that arises when there is a single case report resulting in a chemical being attributed sensitizer status on merely circumstantial evidence. Over a third (31) of the asthmagens described in the study were included on the basis of a single case report. It is likely that a non-trivial proportion of these are misclassified.

With these considerations in mind, the diagnostic criteria for the study do not seem strict - if the author specified a causative agent as causing sensitisation this compound was accepted unless it failed on other criteria (see p50) or unless there were clear errors in the report. However the study has looked specifically at compounds believed to cause occupational sensitisation *in situ* rather than in a laboratory. It is therefore not unreasonable to use a clinical rather than a pathological basis for study acceptance criteria, given that the possibility of compounds being misclassified has been acknowledged throughout the thesis.

The data collected in this study include a number of clinically relevant parameters such as prevalence rates, lowest observed effect concentrations[1], latency and basis of diagnosis. These data are not available for all compounds. Few published reports describe ambient workplace levels of the chemical, those that do are subject to varying degrees of measurement error, and in any case concentrations may differ depending on where the sampler is placed.

Perhaps the most useful parameter to know would be the lowest observed effect concentrations for sensitisation but unfortunately this datum is rarely recorded. The portability of personal samplers must be weighed against their lower sensitivity when compared with 'gold standard' apparatus. The comparability of data from different samplers for

[1]The 'lowest observed effect concentration' is a toxicological term for a threshold concentration above which toxicological effects may be observed.

different chemicals under different work environments introduces too many uncontrolled variables. The presence of confounding chemicals in the work environment and the potential for synergistic mechanisms makes accurately identifying the true hazards difficult . For this reason the study has concentrated on identifying chemical sensitisation hazard rather than trying to assess risk.

The collection of data that was not directly used in the resulting modelling of sensitisation hazard may appear to have been unnecessary. It was not known however whether there would have been enough of these observations to be analysed within the logistic regression model. Some of the data, such as the immune profile and the type of asthmatic response data proved useful for sorting the asthmagens into categories of particular interest. Other data such as the 'prevalence' and the 'method of diagnosis' details were used to indicate that the model performed better with compounds that were unlikely to have been misclassified (see Figures 3.5 and 3.6).

The Role of IgE

Several interesting observations arose from the immune profile data of the active compounds. Evidence for an IgE-mediated response was not found with all chemicals (see Table 3.4) suggesting that mechanisms independent of IgE may be capable of producing asthma[1]. The lack of a statistically significant difference between the molecular masses of those asthmagens for which specific IgE was identified and those in which IgE appeared absent (see Table 3.4) suggests that mass is not the sole determinant of whether an asthmagen elicits an IgE-dependent immune response. In fact, even quite low molecular mass molecules (such as phthalic anhydride, mass$\approx$148) can elicit IgE-dependent immune responses by acting as haptens [16, 57].

[1]This assumes that asthma is defined in terms of disease symptoms rather than in terms of a single immunological mechanism.

Asthma caused by exposure to low-molecular mass organic compounds also displays a heterogeneity of presentation. In particular, clinically defined OA does not always clearly fit into a classical type 1, IgE-mediated immune response [170]. Whilst compound classes such as the acid anhydrides clearly do fit into this mechanistic class, others such as asthma induced by formaldehyde and ethylene diamine do not. Furthermore, studies of the most common cause of OA, toluene di-isocyanate (TDI), are notable for the lack of consensus with respect to the underlying mechanism. In a review Butcher and Salvaggio state that for TDI asthma "<20% of individuals, proved by inhalation challenge to be reactive to TDI have IgE antibodies demonstrable by RAST[1]". This may be an indication that OA may also be caused by a mechanism independent of IgE, a possibility that has implications for the use of predictive systems that presume asthma is caused solely by and IgE mechanism [171].

Asthmatic Response Type

The type of asthmatic response varied (see Table 3.5) between chemicals (and indeed between patients sensitized to the same chemical). The five compounds[2] for which all three response types were reported can be considered highly reactive compounds. It is possible that these compounds could produce asthma like symptoms by several different methods. The isocyanates are a recognised reactive class of chemicals. Azodicarbonamide has two carbons which are highly susceptible to nucleophilic attack (due to the planar the electron withdrawing effect of two nitrogens and one oxygen on each).

One of the unusual aspects of asthma to some low molecular weight compounds is that there can be a delay of several hours between res-

[1]RAST - Radioallergosorbent Test.

[2]azodicarbonamide; hexamethylene di-isocyanate; toluene di-isocyanate; plicatic acid and a trimer of hexamethylene di-isocyanate

piratory challenge with an asthmagen and the resulting asthmatic response. The occurrence of a late asthmatic response was noted with approximately half the asthmagens. This has implications for the identification of low molecular weight chemically induced asthma - the delay between exposure and response can confound attempts to identify the causal agent. In particular it is possible that many cases of asthma may fail to be described as occupational - thus under-reporting of cases may occur.

When the asthmatic response data and the immune profile data were looked at together it was evident that all compounds for which IgE was noted also exhibited a dual asthmatic response in at least one of these cases. In addition, no compounds for which specific IgE was observed caused solely a late asthmatic response. This opens up the possibility that asthma to chemicals that do not cause a dual response may be invoking an IgE-independent means of sensitisation. It follows that the observation of a dual asthmatic response may strengthen the case for suggesting a compound sensitizes via a Type 1, IgE-dependent mechanism. Acid anhydrides and reactive dyes are typical of compounds that fall in to this latter category.

Both acid anhydrides and reactive dyes would appear to be excellent examples of compounds which produce asthma via a Type 1 immune mechanism. It is of interest to note that diisocyanates, the most common and potent asthmagens, seem to less readily fall into this category, perhaps an indication of either a different mechanism or mechanisms. Serum IgE has however been demonstrated to be elevated in mice exposed to toluene di-isocyanate cutaneously[17].

Accompanying Symptoms

Over half the sensitising compounds had nasal symptoms as an accompaniment of OA. Wheeze and cough were similarly widely reported.

Epidemiological studies have shown that rhinitis in particular is more common than asthma [172]. The occurrence of these symptoms amongst workers exposed to known occupational asthmagens may be the first signs of OA developing.

4.2 Speciation and Structure Determination

The chemical criteria for the study - that compounds should be of molecular weight < 1000 and be carbon containing - were chosen to avoid protein allergens (e.g. papain [173, p375]) and transition metal sensitisers (e.g. platinum[173, p468]). Although identification of Chemical Abstracts (CAS) Registry Number was desirable it was not essential. It was more important to identify the correct chemical structure. Often this identification process proceeded via a CAS Registry Number but far too often the chemical described in a paper would refer to an ambiguous trivial name.

Identifying the correct structure proved difficult for several compounds as an ambiguous trivial name was given but a structure was not. An example of this was evident in papers using the term 'alkyl' within a structural name. Consequently the structures alkyl cyanoacrylate and benzalkonium chloride remain ambiguous. However properly speciated representatives of these structure classes (ethyl cyanoacrylate and lauryl dimethyl benzyl ammonium chloride) were reported elsewhere and thus were included in the study.

Six potential asthmagens were excluded because either a unique structure could not be identified, or the true active component of a mix was not identified, or because the compound was missed in the initial search. These are listed in Table D.2. A full list of reported asthmagens (including those rejected) can be found in Appendix C.1. It is of note that the three isomers of methyltetrahydrophthalic anhydride

(MTHPA) (which were excluded because the true asthmagen amongst them could not be identified) have the same predicted asthma hazard (0.93). This is due to the nature of the predictor fragments in the model not differentiating between the structures. It is possible all three isomers can cause asthma. It is worth noting that had MTHPA been included in the study the HOR for the acid anhydride fragment would have been of even greater magnitude.

The constituents of the substance EPO 60 could not be used in the analyses because the true asthmagen amongst them was not identified (see Table D.2). The nature of the compounds and their subsequent predicted hazards indicates that two of them, tetraethylene pentamine and isophorone diamine have a high hazard prediction, suggesting either (or both) as the possible cause. (The relatively low hazard prediction for 4,4'-diamino-diphenyl methane suggests that this is less likely the cause although it may be considered similar chemically to the reported asthmagen paraphenylene diamine (see Figure 4.1). However the predicted hazard was also low (0.17) and the original case report [100] on whether paraphenylene diamine is a true asthmagen may be flawed[1].

4.3 Substructure Searching

In order to relate activity to structure it was necessary to create substructure searching algorithms to identify "fragments" contained within chemical structure. The structure of each compound in the study was described in terms of fragment content . To achieve this the tabulation

[1]The diagnosis is based on patch testing of patients which merely indicates Type IV (delayed) cutaneous rather than respiratory sensitization. In addition the workers were exposed to fur and so their asthma symptoms may have been due to animal dander allergy. Finally, the results in asymptomatic asthmatic patients of positive inhalation challenge tests to paraphenylene diamine and hydrogen peroxide may have been due to the increased airway hypersensitivity exhibited by asymptomatic asthmatics.

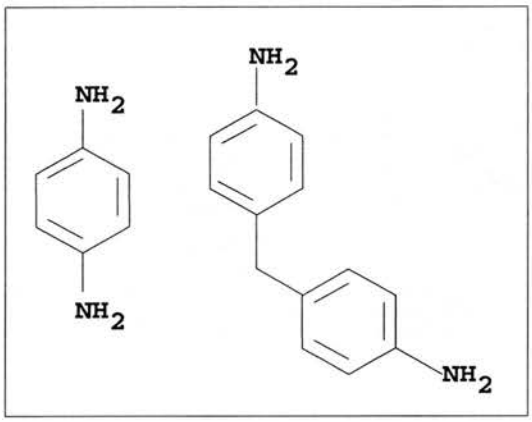


Figure 4.1: The structures of paraphenylene diamine and 4,4'-diaminodiphenyl methane.

of fragment content was automated computer programs (see Appendices A and B). Although commercial software was available that provided this functionality, the restrictive nature of commercial software would have prevented the creation of a stand-alone hazard prediction program. Since the creation of a practical hazard prediction system was intended from the outset (see Section 1.5) the decision was made to write the required software. This brought considerable programming complexity into the study and while it will be dealt with briefly here in the discussion it did account for a considerable amount of the work.

The computer programs could describe compounds in one of two fragment types (see Figure 2.6). The Methods section describes in detail the two types of fragments used (type 1 and type 2 fragments, p62). The selection of fragments was not objective and consequently this may be seen as a source of bias. Whilst the use of a validation dataset helps to counter such a criticism it would have been preferable to have had a systematic (thus objective) method for generating all possible fragments[1] for the dataset. It should be noted that even if implemented,

[1]Although an additional requirement would be that the fragments would have to occur a minimum number of times in the dataset to be of statistical importance.

automatic fragment generation would probably only occur for Type 1 fragments.

The main advantage of type 1 fragments is the speed with which new fragments can be added. The fragment is simply drawn in a chemical drawing program and a `*.mol` connection table file [159] (see Figure-2.4) created. A fragment dictionary of compounds can be entered into a database and from this a `*.sdf` file [159] (see Figure 2.5) created. The disadvantages of type 1 fragments are that searching for them is computationally intensive and that they do not allow 'wild-card atom or bonds to be specified.

Type 2 fragments are useful for the same reasons. A particular atom node in the fragment can be specified to match any one of a list of elements and similarly a bond may be allowed to match irrespective of whether it is saturated or not. Furthermore, because type 2 fragments are hard coded into the program (as subroutines), the search algorithm can be optimised to match against the least frequently occurring elements of the fragment first. In this way fewer comparisons are made before a match can ruled out. However because type 2 fragments are hard coded, adding new fragments is technically demanding; this is the main disadvantage of type 2 fragments.

The coding of the type 2 fragment algorithms chronologically preceded the coding of the type 1 fragments. This occurred because the algorithm to implement a general (type 1) substructure search is far more complex and fits into the set of problems known as *NP* complete [174]. Consequently the algorithm involves exhaustive and recursive searching to find the result. The complexity of the algorithm can be guessed from attempting to formalise the steps one would take to determine whether one compound is contained within another. In order to do this, one must match every atom of one compound against every atom of the other[1]. If an atom matches then all its neighbouring atoms need to be

[1]There are some screening algorithms which can be used to avoid some of the

matched. In addition, the bond type between them needs to be checked. All these checks have to be repeated as the matching progresses from bond to bond (and back).

There are further difficulties. There is a need to know when to stop: for cyclical searches checks need to be made to prevent infinite looping around a ring. In addition many structures have symmetry and so checks need to be made to ensure that matches are not occurring on the same atoms in reverse. The solution to this was to identify how many times a fragment could find itself in itself then correct the result accordingly.

4.4 Previous Structure-Activity Studies Of Chemical Asthma Hazard

To date most methods for predicting chemical asthma hazard have been based on animal models and a number of methods exist *in vivo* to screen compounds for OA hazard [175]. Several animal models have been developed in the last few years, perhaps the most promising of which is the use of a cytokine fingerprinting technique in the mouse [176]. These methods however tend to be expensive, time consuming and specific to only one or two classes of compound. There was a clear need for a cheaper, more rapid method of identifying potential respiratory sensitizers prior to their being introduced into the workplace.

Current quantitative structure-activity approaches to investigating the relationships of respiratory sensitisation have been limited to using small datasets of sensitizers [177]. A qualitative chemical structure-based hazard prediction exists for sensitising compounds in the form of a rule-based, expert system (DEREK)[178]. The use of systems such

exhaustive searching.

as DEREK relies on an "expert" judgement including mechanistic assumptions and the determination of empirical rules.

More recently Karol *et al.* published a structure-based model for predicting skin sensitisation hazard which does not assume a mechanism of activity [179]. Only 39 "respiratory chemical allergens", based on animal studies as well as human data, were used in the study. These case chemicals were compared with compounds which did not cause skin sensitivity.

The identification and selection of the active and control compounds for all studies may affect the validity of the results. Our selection of controls was based on agents that are known or presumed to be hazardous and to which there is significant occupational exposure to warrant the setting of exposure limits. It is therefore probably more appropriate than random selection of chemicals (regardless of human exposure) or assuming (as Karol *et al.* did [177]) that skin and respiratory sensitisation are mutually exclusive and that skin sensitizers represent adequate controls for respiratory sensitisation. In particular the requirement that an active compound be a documented cause of human occupational respiratory sensitisation rather than suspected from screens *in vivo* was fundamental to the study. The use of chemical cases defined on the basis of animal models may introduce a bias in favour of a particular pathological mechanism.

4.5 Clustering

Clustering was the first method used to differentiate between compounds of differing activities by means of their structural information. It provided a useful learning tool but it was dismissed as a suitable modelling option early in the study. For this reason there is no cluster analysis of the differences between skin and respiratory sensitizers. This section will look briefly at the role of clustering

Clustering techniques involve the classification of *entities* by *descriptors* to produce groupings based on similarity. Several methods can be employed and these fall in to two distinct categories: hierarchical; and non-hierarchical. The method chosen in this study was the divisive hierarchical Guenoche method [166] based on the results of a preliminary study performed by Dr G. Downs. This preliminary study looked at several other clustering methods but as a cluster program had to be written from the general algorithm (see p70) it was decided to settle on one clustering method. The Guenoche method was chosen after considering the results of Table 3.7, and following discussion with Dr Downs.

Clustering does not directly provide a method of predicting a biological activity for chemicals. However, an activity for a compound may be inferred by association. That is, it is inferred that the activity of a compound will be similar to that of the other compounds with which it shares a cluster group. Calculating a chi-squared statistic for the clusters based on this method results in very significant p values as the number of clusters is increased. These p values in isolation are however a poor and inappropriate measure of overall model performance. The results of the preliminary cluster analysis are summarised in Table 3.7.

For the complete asthmagen and control dataset clustering was performed using either 'presence or absence' of descriptors - the binary method, or using the occurrence frequencies of the descriptors - continuous method (see p69). A schematic representation of clustering process of each is given in Figures 3.2 and 3.3. Several of the clusters appeared to be strongly predictive of either control or asthmagen status. Cluster 5 of the binary method was one such cluster. All 31 of the compounds in cluster 5 were controls and 29 of these contained at least one halide (fluorine, chlorine or bromine). This is worth noting as the possibility that halide may confer a 'protection' against a compound being an asthma hazard will be mentioned again.

Despite a few interesting clusters, overall the clustering approach proved unsatisfactory for a number of reasons. Firstly, it is not easy to identify the important discriminating variables nor quantify their contribution, so little mechanistic insight can be gained from this method. In addition, the use of excess (redundant) descriptors may occur and this is to be avoided [158]. The use of such redundant descriptors increases the possibility that, for the learning set, the differentiation will occur by chance. The descriptors will work for the learning set (by chance) however when the same method is applied to a novel set the descriptors will not be applicable. Increasing the number of clusters may improve kappa but at a cost. The resulting smaller clusters are more likely to contain solely asthmagens or solely controls however the smaller size may preclude a cluster being statistically significant. Once clusters with very few or even single entities start to form the value of clustering for activity inference becomes suspect and the kappa value meaningless. It would be very simple to use cluster analysis to produce N clusters were N is also the number of entities used. Such a procedure would produce a prediction system that was perfect for the learning set but incapable of providing reliable predictions for new compounds.

Consequently clustering is perhaps a poor choice for predictive modelling but it may still have a role for classification of molecules into mechanistic groupings. It was not however considered further in this study.

4.6 Hazard Odds Ratios

The Hazard Odds Ratio (HOR), in providing a relatively simple measure of a fragment's contribution to activity, also provides an excellent fragment screen prior to use in multivariate analyses (such as regression). In any multivariate study it is desirable to maintain a sufficiently greater number of observations than variables to avoid vari-

able redundancy [158]. For QSAR[1] were the scope exists for the use of large numbers of variables (from fragment dictionaries), HOR's provide a means to individually screen out those variables most likely to be redundant. This approach was used to great effect with the type 1 fragments (see p2.3) and resulted in the identification of a number of potentially important fragment variables (see Table 3.8).

One of the most interesting outcomes of this study is that the HOR's for some fragments, for example aliphatic amine groups, markedly increase with occurrence frequency (see Table 3.11). Furthermore the HOR's for single occurrence of chemical substructures such as imine fragments, aliphatic or aromatic amines or carbonyls were not significant at the 5% level. However two or more occurrences of these fragments was significant. It appears that a single occurrence of certain hazard fragments in a compound may not be sufficient to pose an asthma hazard. Instead some degree of bi- or poly-functionality appears to be required. These results support the hypothesis that chemicals that are bi-or poly-functional (i.e. have cross-linking) are over represented within the set of known respiratory sensitizers. Such an hypothesis implies a threshold of two active groups, and that a single fragment is associated with low hazard. The results could however be explained by arguing that whilst a single fragment was insufficient to produce a statistically significant HOR it is still hazardous but only to a fraction of the hazard of two fragments.

Whilst the evidence for a cross-linking mechanism in OA is not totally conclusive it should be noted that no reports of isocyanates causing asthma are due to mono-isocyanates. This could however be a reflection of infrequent use since bi-functional compounds find wide industrial application precisely because of their cross-linking properties. However one of the notable features of isocyanate asthma is the fact that many studies fail to identify evidence for an IgE-mediated sensitisation mechanism [8, 85, 180] in every patient. This is not to say it has not

[1]Quantitative structure-activity relationships.

been shown in some cases, rather, that it is frequently absent despite the presentation of OA. If a cross-linking mechanism is responsible for some asthma cases, it may be for those atypical cases where IgE has not been demonstrated. Indeed, the implication is that di-isocyanates may be able to cause asthma by two distinct mechanisms.

There are several other compounds which can be viewed as potential cross-linkers. Although formaldehyde ($H_2C=O$) has a structure which appears mono-functional, in aqueous environments it acts as a diol ($HO-CH_2-OH$). One of the uses of formaldehyde is as a fixative and achieves this by cross-linking. No IgE has been determined in cases of formaldehyde asthma. Similarly, glutaraldehyde (which is a bi-functional aldehyde) is also a cross-linking agent. Again no IgE has been determined. Neither compound gave rise to an immediate asthmatic response type which would be expected with a Type I IgE-mediated asthma. Interestingly the carbonyl group is one of those for which two fragments are significantly hazardous whereas one is not (see Table 3.11).

Amines also seem to be far more hazardous in pairs (see Table 3.11). The bi-functional chemical ethylene diamine is an asthmagen yet the mono-functional ethyl amine is not. Furthermore no IgE has been demonstrated for ethylene diamine induced asthma and the response type in a case where bronchial provocation testing was performed was 'late dual' [181]. Another curious feature of ethylene diamine is that it is better known as a skin sensitizer [182].

A mechanistic hypothesis based on the need for chemical bi-functionality is not inconsistent with immuno-biochemistry of antibody-allergen recognition. Theoretically, immunologically recognisable epitopes may arise from a normally non-antigenic proteins if it is cross-linked. Whether this is important through intra- or inter-molecular cross-linking requires further investigation. However, amongst the compounds for which cross-linking may be important, are the poly-amines and the polyisocyanates which can cause sensitisation without any observed

IgE response. This contrasts with the more classical hapten-induced, IgE-mediated response seen with the anhydrides. Yet the sensitisation due to these compounds is certainly an allergic response. The possibility remains that bi-functional sensitizers may sensitize via a non-classical IgE mechanism such as by causing membrane damage or cross-linking of a cell surface receptor. A possible candidate cell would be the mast cell. Indeed, to activate mast cells, IgE 'cross-binds' two cell surface receptors. These seem unlikely hypotheses at present since they need to be reconciled with the allergic-like nature of the disease and the observation that certain people are more susceptible than others. However if airway insult by irritant gases or disease is important as an initiation process in asthma pathogenesis, then a sensitisation may be the result of chemicals reaching the normally unexposed airway sub-epithelial cells.

Anhydride Asthma

One of the few fragments for which a single occurrence is significantly associated with reported OA hazard is the acid anhydride fragment. Although acid anhydrides contain two carbonyl groups and in this respect may be considered bi-functional, these are not independent. When one reacts, perhaps to form an amide with a lysine (that is, an amine side chain) residue of a protein, the remaining carbonyl group becomes part of a carboxylic acid. Carboxylic acids tend to be more chemically stable, so although these groups could form esters with other proteins this is unlikely. Published evidence from elsewhere has suggested that acid anhydride asthma is caused by the anhydrides acting as haptens and rather than cross-linkers [183, 57]. Anhydride asthma is typical of Type 1 (immediate hyper-sensitivity) IgE-mediated asthma.

One of the arguments used to suggest that anhydrides work by cross-linking is the fact that acetic anhydride is not an asthmagen[1] because

[1]According to the SWORD scheme acetic anhydride has been reported as an asth-

it cannot bridge - when it undergoes hydrolysis the molecule splits into two acetic acid molecules. This argument is weak. It could equally be argued that an acetyl group makes a poor epitope when compared with the phenyl group of phthalic anhydride. So acetic anhydride may be reactive enough to attach to proteins, but in doing so does not make the host protein antigenic. Therefore it is unlikely that cross-linking and bi-functionality have a role to play in the pathogenesis of anhydride-mediated OA.

Protective Fragments

The use of substructure hazard odds ratios may also identify 'protective' substructures, that is, chemical substructures which may preclude a chemical from presenting an OA hazard. This study showed that the presence of aliphatic chlorine in a structure was associated with a reduced likelihood of that structure being a reported cause of OA. This may be due to increased toxicity[1] of chloro-compounds (such that other adverse effects occur at lower exposures than required for sensitisation).

Odds Ratios To Differentiate Skin and Respiratory Sensitizers

When odds ratios were used to differentiate between asthmagens and skin sensitizers the carboxylic acid fragment appeared to be associated with a compound being an asthmagen. The carboxylic acid fragment had a large and highly significant trend favouring its presence in asthmagens (see Table 3.13). The question of what role the carboxylic acid group plays in asthmagens arises again. However the reason for

magen however all SWORD reports should be considered as requiring confirmation. Acetic acid however has been reported in published literature as causing asthma [184].

[1]That is toxicity other than asthma toxicity.

the high significance could be that the fragment is excluded or rare amongst skin sensitizers because the carboxylic acid is a hydrophilic group. Skin sensitizers are typically more hydrophobic than respiratory sensitizers [185]. This would explain why the fragment did not appear so significant in the differentiation of asthmagens from the control dataset (which included skin sensitisers, see Figure 2.1).

The presence of oxygen but not nitrogen confers an increased likelihood that a compound is an asthmagen rather than a skin sensitizer (see Table 3.12) which may be mostly due to the acid anhydride and the carboxylic acid fragments having large HOR's. When these fragments are taken into account, there is probably a fairly even spread of oxygen and nitrogen compounds across both skin and respiratory sensitizers. Perhaps one of the factors that determines whether a compound is a skin or respiratory sensitizer (or both) is not how it reacts but where it reacts, and this may be a function of its water solubility. Given that skin is generally a dry environment whereas the respiratory tract is moist this might seem a reasonable hypothesis. Polar compounds on the skin will not be readily absorbed and are easily washed away. Lipid soluble compounds however will be quickly absorbed and are difficult to wash away. In contrast, the moist environment of the airways favours the polar compounds.

4.7 Logistic Regression

The principle advantages of using logistic regression are that variables can be assessed for a statistically significant contribution before inclusion and and that allowance is made for variable interaction. Thus the contribution made by fragments may be different when they are considered in combination as opposed to in isolation.

The use of the stepwise logistic regression facility within the SPSS program was considered appropriate since this work falls into both the

exploratory and the predictive categories described by Menard [186]. Briefly, all the fragments entered into the logistic regression modelling program start in the model. Variables are then removed on the basis that their contribution to the model is not statistically significant. The significance of the likelihood ratio was used in preference to the Wald statistic to determine whether a fragment stayed in the model because some of the fragments, for example the isocyanate fragment were rare but occurred exclusively in active compounds. That is, the presence of the fragment guarantees activity but its absence does not preclude activity.

The first logistic regression model (see Table 3.14) distinguished between asthmagens and controls with a kappa (based on the validation) of 0.74. When internally validated (using the learning dataset), the logistic regression model appeared to be better at predicting asthma hazard for those compounds for which the confidence in their being asthmagens was greatest. Compounds for which diagnosis was based on respiratory challenge (the definitive diagnostic method) (see Figure 3.5) or those compounds for which two or more cases of asthma had been reported were in general better predicted (see Figure 3.6).

Analysis of the misclassified compounds highlights some key points about the model. It is possible that if we assume the model is a good predictor of hazard, the misclassified compounds represent data outliers. Consequently, by looking at misclassified active compounds such as styrene, ethylene oxide and chloroxylenol we may be looking at asthma caused by an atypical mechanism. Equally, for the misclassified controls, such as methacrylic acid the model might be correctly identifying compounds with true sensitising hazard which have not yet been reported as such. For example the control compound methacrylic acid is predicted as a hazard (0.77) and it should be noted that methyl methacrylate is an accepted respiratory sensitizer and acrylic acid is a suspected one [108]. Another curious false positive was strychnine (0.74). It can be postulated that strychnine may be a potential sensi-

tizer but because of its reputation as a poison, exposure to it may have been more tightly controlled than compounds of equivalent but less notorious toxicity[1].

The correlation between hazard predicted and confidence of original active classification, poses the question of whether it would have been valid to classify compounds by potency or prevalence for use in a linear regression analysis. There are good reasons why this would be a less effective approach to modelling. The first reason to avoid using activity coded as a continuous variable is that the data on which to base this value is simply not available for all compounds. Secondly, the 'basis of diagnosis' and the 'absolute prevalence' may be measures of the confidence with which asthma causality can be assigned - the true value is either 'yes' or 'no' - but combining the two into a meaningful scale would require a great deal of subjectivity. It would be difficult to decide how many separate cases of sensitisation diagnosed on circumstantial evidence are equivalent of one case diagnosed with a double blind respiratory challenge.

There are several notable substructure fragments absent from the logistic regression model variables (see Table 3.14) which had significant HOR's (see Table 3.10). For example neither the carboxylic acid fragment nor the ethanolamine backbone are represented in the logistic regression model variables (Table 3.14). This suggests that the HOR's for these fragments occurred because these fragments contained other substructure fragments, such as oxygen atoms or nitrogen atoms. When corrected for the presence of these smaller fragments the ethanol amine fragment would no longer appear independently significant.

The most striking feature about the fragments in the first logistic regression model (see Table 3.14) is the magnitude of the aldehyde and ketone fragments. The aldehyde fragment has a large negative contribution (-12.21) which may seem unexpected given that formaldehyde

[1]Not asthma toxicity.

and glutaraldehyde are amongst the asthmagens. The standard errors for these two fragments are also large, indicating that these fragments are may have been included by chance. However the *p=0.0043* value of log likelihood ratio for the ketone fragment suggests it has an important role in the model. The definition of the ketone fragment does not preclude it matching the same carbonyl group as the aldehyde fragment. It is notable that the magnitude of the aldehyde and ketone β coefficients (-12.21 and 14.36 respectively) are of similar magnitude but working in opposition. It is probable that the important carbonyl substructure in determining asthma hazard is a carbonyl fragment that is *not* an aldehyde.

The acid anhydride, isocyanate and acrylate groups have large positive β coefficients and small standard errors (relative to their β coefficients). This suggests that these compounds not only influence whether a compound is hazardous but may also be decisive. That is, their presence alone is enough to determine whether they are predicted active or not.

The model also feature three halide atoms, chlorine (aliphatic and aromatic), bromine and fluorine. These all appear to have a 'protective' effect with respect to asthma hazard. This 'protective role' has been also been noted with the clustering and the HOR's. It may be appropriate in future work to combine these fragments into a single generic halide type 2 fragment.

The model might find application after further validation for the purpose of hazard prediction - thus if a novel chemical entity were to score a value of say higher than 0.3 or 0.5 it may be cause for concern if it were not being used in a manner in which exposure was stringently controlled and prospective health surveillance practised. Caution should be advised however with respect to identifying 'safe' compounds. It is apparent that the hazard posed by some known (but now less commonly reported) respiratory sensitizers such as para-phenylene diamine are poorly predicted by the model.

The Model to Differentiate Between Skin and Respiratory Sensitizers

The role of the skin dataset has been throughout this study been relegated to a lesser importance than the task of identifying respiratory sensitizing agents. This is partly due to the difficulties of finding suitable controls for skin sensitizers. However the role of the skin dataset was always intended to allow comparisons between skin and respiratory sensitizers, rather than to attempt to predict skin sensitizing hazard.

A number of compounds cause both types of sensitisation. When a prediction is calculated for these compounds by the skin model they appear across the full range between skin and respiratory ends of the scale (see Figure 3.7). It has been suggested that skin and respiratory compounds are functionally divergent [56] and perhaps mutually exclusive. Whilst this may be the case in laboratory studies, the situation in the workplace is more complex. Models of sensitisation which are based on a divergence of skin and respiratory sensitisation may fail to identify hazardous chemicals because they presume a single mechanism of sensitisation.

4.8 Validation

The validation of the model by an independent set of chemicals is a key factor in evaluating this work. It alone is sufficient to make this work unique amongst structure-activity studies in this field. Similar studies such as those of Karol *et al.* [177] and Payne *et al.* [178] have failed to validate their systems independently.

The validation of the model utilised two sources of reported respiratory sensitisers. The first and perhaps less reliable source (due to the ab-

sence of a quality-control element) was the reports to SWORD[1] [163]. The second source were the data published after 1994. This second source was probably more reliable since a usual prerequisite of publication is that the report be peer-reviewed, thus reducing the the possibility of speculative speciation of compounds as asthmagens. This is a possible explanation of why the model performed better when validated by published data rather than SWORD data. A possible source of bias that would effect the kappa values for the SWORD data is the fact that several of these chemicals were originally used as controls. Chemicals 'promoted' from control to SWORD status are subject to a bias that results from their originally contributing to the model definition of 'control'. Allowance was made for this fact (see Figure 3.10 and Table 3.20) and it was very apparent that predictions were (not surprisingly) poorer for those SWORD compounds which had been controls in the original learning dataset. The poor fit of some of the SWORD data to the model is a cause for concern since it suggests that either the model doesn't work or that the SWORD scheme produces unreliable data.

The nature of a scheme such as SWORD is such that false positives are less of a concern than false negatives. One possible false positive in the SWORD scheme list of suspected asthmagens is halothane (hazard index 0.00). Personal communication with an experienced anaesthetist and an absence of published reported cases of halothane induced-asthma despite exposure under a clinical setting of both patients (high concentrations) and anaesthetists (low but frequent concentrations), suggests halothane may be such a false positive. (It is however of interest to note that another inhalation anaesthetic, enflurane was found to have been reported to cause OA in 1976 [131]).

Further limitations of the SWORD Scheme exist. Some compounds in

[1]Surveillance of Work-related and Occupational Respiratory Disease, a scheme introduced in 1992 to allow chest physicians to report occupational groups and agents with a high risk of respiratory disease (including asthma).

the SWORD list appear twice: methylene *chloride* and methylene *chlorine* probably refer to the same compound; likewise, *Trichlorethane* (as spelled) and *trichloroethylene* also seem to refer to the same compound. The verification of a compound's asthmagenicity by a suitable respiratory challenge test is not required for a compound to be included in the SWORD Scheme. Furthermore the unique identification of a single chemical cause is often missing with the result that the suspect agent is not correctly speciated.

Despite the limitations of the validation data the model performed well, particularly for the peer reviewed data where the kappa value of 0.81 is extremely satisfactory (see Table 3.20). Although this is based on relatively few validation compounds it is nonetheless comparable if not better than existing methods which have not been as rigorously validated. Consequently the Internet based prediction system that has been developed directly from the programs used and the logistic regression models created provides the first readily available prediction systems for chemical asthma hazard.

If discrete judgements on activity are to be made a suitable threshold is required. Normally a threshold of 0.5 is used but the data here exhibits a kappa value consistently around or above 0.7 (see Table 3.20) for thresholds in the range 0.2 to 0.6. One reason why a value of below 0.5 is more appropriate in this study is because the model is based on data using a considerably greater number of controls than asthmagens. Perhaps a more convincing reason for choosing a lower threshold is that there is less to be lost in predicting a compound active when it is not than missing a true sensitizer. Taking this into consideration, the Internet based system considers a predicted hazard of 0.3 or above as cause for concern.

Postscript

Several compounds not included in the validation have recently been reported to cause occupational hazard. They were identified via a Med-Line search so only the title and abstract has so far been consulted. However because testing of the a compound takes minutes and costs very little, hazard predictions can generated with ease. The model predictions for these compounds are given below:

Compound	Hazard
Cefmetazole [Fracchia *et al.*, 1996]	1.000
Ninhydrin [Piirila *et al.*, 1997]	0.045
Isothiazolinone [Bourke *et al.*, 1997]	0.123
Diethanol-amine [Piipari et al., 1998]	0.957
1,2-benzisothiazol-3-one [Moscato et al., 1997]	0.605
Morphine [Ulinski *et al.*, 1996]	0.366

Of these the model would have identified cefmetazole, diethanol-amine and 1,2-benzisothiazol-3-one as very hazardous compounds. Morphine would be considered a hazard as it is greater than 0.3. It has predicted poorly for the other two compounds but this is not necessarily a problem. Since the methodology has been validated, the model could be updated to use newly diagnosed asthmagens.

4.9 Future Work

This work provides a solid foundation for several follow up studies. Firstly there is now a case for updating the predictive model to include those occupational asthmagens identified (from case reports in peer reviewed journals) since 1994. This would provide the most up to date basis for assessing novel chemical structures for asthma hazard.

The selection of fragments for the logistic regression could be improved. For example the creation of a generic halide type 2 fragment would pool the chlorine, fluorine and bromine (and possibly iodine) atoms to see if the protective effect is general across the Group 7 elements.

A second way in which the model could be improved is to validate it against work currently being undertaken to assess chemical asthma hazard by *in vivo* assays [171]. This would provide an interesting cross-validation of two techniques. Any agreement between the two methods would provide strong evidence for the acceptance of both the structure-based prediction system and the mechanisms underpinning the laboratory work. Conversely, any areas of disagreement would highlight areas were laboratory animal-based studies cannot be reconciled with the available human data.

The data used for this study may also be suitable for the creation of a neural network based predictive model. One drawback of using neural nets is the difficulty faced in identifying the relevant data input. This precludes (or at least makes very difficult) the inference of underlying chemical mechanistic hypotheses.

A more long term goal might be to apply the methods used here in a slightly different area. It may be possible to apply the technique successfully to identify carcinogens by structure alone in a more general way than has previously been attempted.

Chapter 5

Conclusions

It is evident from the reviewed literature that clinically defined occupational asthma cases may arise in the absence of any evidence of an allergic mechanism. That is, asthma occurs without an identifiable immunological pathway being invoked. Consequently a distinction needs to be made between 'occupational allergic disease hazard' and 'occupational asthma hazard'. The former implies an immunological process whereas the latter implies merely that a process leading to asthma[1] is involved.

It is evident therefore that the original null hypothesis (see p15) which states:

> The potential of a chemical to cause occupational allergic disease cannot be predicted from structure alone.

was overly optimistic. This work does not provide a basis for rejecting this null hypothesis. Had the null hypothesis been more conservative such that it stated *'The occupational asthma hazard posed by*

[1]Defined clinically by symptoms.

a chemical cannot be predicted from structure alone' then there would be grounds to reject it. Unfortunately null hypotheses cannot be made retrospectively!

Despite failing to reject the null hypothesis this work provides much that is useful such as (1) the novel methodology, (2) the general observations about the types of chemicals that pose an asthma hazard, (3) the comparison between skin and respiratory sensitisers and (4) the production of a prototype hazard prediction system.

1) The chemical case-control methodology is an effective and novel approach to quantitative structure-activity relationships. The decision to pursue an observational rather than a mechanistic approach was crucial factor in the validity of this work. The results are independent of any *a priori* assumptions of mechanism and are based on a statistical approach. The results can therefore be understood in the absence of immunological mechanistic knowledge. Furthermore the results are better than those obtained in other studies.

2.a) Cross-linking may play a key role in an IgE- independent asthma mechanism. A number of chemical groups such as amines, aldehydes and isocyanates, when present twice or more in a chemical, are sufficient to make that chemical an asthma hazard. The target of the cross-linking is not known but mast cell IgE receptors are likely candidates.

b) Acid anhydrides probably do not cause sensitisation by cross-linking. Acid anhydride-induced asthma is typical of a Type 1 immediate hypersensitivity asthma and as such probably causes asthma by an IgE mediated response rather than by a cross-linking mechanism.

c) The presence of halogens (fluorine, chlorine and bromine) in a compound is associated with a reduced asthma hazard. This may be a result of other toxic effects preventing workers being exposed to sensitizing concentrations.

3.a) Skin and respiratory sensitizers are not mutually exclusive. Although differences exist between skin and respiratory sensitizers there is considerable overlap. The same compound may cause both types of sensitisation. The functional differences may arise from differing physical properties, particularly water solubility. Compounds causing occupational asthma tend to contain more hydrophilic fragments.

4) *A simple, working prototype hazard prediction system has been created.* As a working hazard prediction model, the findings of this work can be quickly, cheaply and easily compared with those of other studies and of emerging case reports. Therefore there is a ready mechanism for extending the validation of the predictive model presented here and, as required, improving upon it.

There are two major weaknesses of this work which need to be acknowledged here and remedied in future studies: (1) the subjective selection of descriptor fragments; and (2) the lack of 3-dimensional information in the description of chemicals.

1) *The non-objective selection of fragments may have biased the results.* Objectively selected fragments were not used because of the difficulty in generating the exhaustive list of possible fragments. The choice of descriptor fragments may therefore have been biased towards fragments known to be present in compounds that pose an asthma hazard. The successful validation procedure suggests that this does not substantially invalidate the findings however this bias must be considered in the overall interpretation of the results. If the difficulty in generating an objective set of fragment descriptors were to be overcome, then the use of such a fragment set should provide a more robust methodology.

2) The descriptor fragments used were too small to contain topographical (3D) chemical structure information. This study is notable for not employing 3D structure information in the chemical descriptions.

Molecular shape is important for many if not most biochemical interactions, consequently the use of topological descriptors in this study may be a cause for concern. The immunological mechanisms involved in antigen recognition are highly dependent upon shape as well as charge and hydrophobicity so one might expect that descriptors discriminating between sensitisers and non-sensitisers would require to have recognisable shape properties. However the descriptors (small sub-structure fragments) used in this study may be too small to influence the overall shape of the molecules from which they are derived. If the descriptors are considered too small to influence overall shape or form, then the implications are either (i) that the results described in this work arise by chance, or (ii) that shape is not the major determining factor of the biological endpoint described in this work. If the former (i) is the case then this work is relegated to being merley a data collecting and methodology developing body of work. If the latter (ii) is the case then it follows that what this work describes is not allergic respiratory sensitisation hazard, but occupational asthma hazard. This is still consistent with the hypothesis that bifunctionality rather than overall form seems to determine hazard. If the compounds are not antigenic in themselves but rather they create antigenic features with host molecules then the descriptors fragments need not contain large amounts of 3D information since they are markers of reactivity and perhaps hydrophilicity/availability not shape.

In many ways the descriptors used in this study are not solely indicators of shape or reactivity. The study relates the presence of chemical descriptors to a process that also involves inhalation and absorbtion as factors influencing the biological endpoint of asthma symptoms. Given that the study models a multi stage-process it is not surprising that the conclusions drawn from the apparently significant fragments are general rather than specific ones.

It is possible that by incorporating chemical structure descriptors based upon shape rather than atomic connectivity better results could be ob-

tained. However there is a limit to the usefulness of such an approach in this study. As the 3D complexity of a descriptor increases, its frequency of occurrence and therefore its potential to be of statistical significance drops.

Overall the conclusion is that the chemical case-control methodology employed in this thesis can be useful if applied to heterogeneous and poorly understood chemical-structure activity relationships. The methodology cannot provide detailed mechanistic insights but may lead to useful general observations. The results and the predictive model provided by this work should be a useful adjunct to other studies of chemical asthma hazard.

Appendix

Appendix A

Computer Programs

In 'C' there is a declaration type termed a `structure`. These '`structure`'s are a method for representing data in an ordered and logical manner. Since an ambiguity between chemical structure and programming `structure` could arise, the latter will always be referred to in `typewriter` font.

`molecule.h`

`molecule.h` is not a subroutine as such but this file contained the `structure` definition for the software coding of molecule connectivity.

Two types of `structure` are defined: the molecule `structure`; and the atom `structure`. It is worth looking at these `structure` definitions closely to understand how chemical structures were represented within the software. The `structure` of an atom is described in Figure A.1

The role of the variables listed in Figure A.1 a is:

`sym` contains a string containing the element symbol.

```
struct *atoms{
  char sym[3];        /* The atomic symbol */
  int  mask;          /* A mask used by some routines */
  int  aromatic;      /* Marked if atom is in aromatic ring */
  int  valence;       /* Default valence for element */
  int  n_bonds;       /* number of atoms bonded to atom */
  int  *bonded_atom;  /* Structure array of n_bonds atoms */
  int  *bond_type;    /* Structure array of n_bonds bond types *
};
```

Figure A.1: 'C' structure representation of an atom.

`mask` contains a masking flag variable used by some routines.

`aromatic` contains a flag indicating whether the atom is part of an aromatic ring.

`valence` holds an integer giving the normal valency for that atom.

`n_bonds` contains the number of atoms this atom was bonded to.

`*bonded_atom` contains an array of `n_bonds` integers corresponding to atom number of the bonded atom in the molecule.

`*bond_type` contains an array of `n_bonds` integers corresponding to the bond types between the atom and the `n_bonds` bonded atoms.

The structure of a molecule is described in Figure A.2. It can be seen that any instance of a molecule contains a variable that holds the location of the molecule name; `*name;` an integer `n_atoms` telling how many atoms are present in the molecule; an array `*atom` of `n_atoms` addresses to the constituent atoms; and a flag integer variable `invalid`.

Using this representation each molecule could be represented in such a way that all the topological information, including bond and atom types could be stored in a manner approaching an intuitive representation of chemical structure.

```
struct molecule{
  int    invalid;          /* Flag variable for some routines */
  char   *name;            /* Molecule name, if one exists */
  int    n_atoms;          /* Number of atoms in molecule */
  struct atoms *atom       /* Array of pointers to each atom */
};
```

Figure A.2: 'C' structure representation of a molecule.

readmol()

This routine reads a molecule in from a standard chemical connection table (MDL *.mol or *.sdf file formats [159] These connection tables contain a list of all atoms followed by a list of all bond pairs between atoms (see Figure 2.4). The readmol() routine interprets an *.sdf file into a 'c' molecule structure (see Figure A.2). These 'c' structures are dynamically allocated memory structures which exist only as long as they are needed.

molfree()

This routine frees the dynamically allocated 'C' molecule and atom structures.

strseek()

This general purpose routine reads forward through a file until it reaches a specified string, which it reads and leaves the current file position as the character immediately following the string.

saturate()

This routine is used to add the implied hydrogens to a structure. In an `*.sdf` file (see Figure 2.4) the hydrogens are not usually explicitly stated but are instead implied by the unused valency. The routine `saturate()` scans molecules previously read in by `readmol()` and recalculates the atom and bond content to explicitly include the hydrogens. This is needed because the same molecule, coded twice with implied hydrogens in one and explicitly stated hydrogens in the other would mean the former could be seen as a substructure of the latter but not vice versa.

findfrag()

The routine `findfrag()` determines whether one structure is a subset of a second. Since some hydrogens might be explicitly labelled in the `*.sdf` file of the first structure which are only implied in the second, the second molecule should first be processed with the `saturate()` routine.

To find if two files contain identical structures, both should be preprocessed with `saturate()`, and `findfrag()` should be run to see if the first molecule is a substructure of the second and if the second is a substructure of the first. If both these calls to `findfrag()` return true then the molecules are identical.

The routine `findfrag()` has a few internal checks to ensure that symmetrical substructure fragments are only found once. Consider ethane as fragment. It can find itself in itself two ways. The problem is multiplied sixfold with benzene - benzene can be superimposed upon to itself 12 ways (not six because it can flip). To resolve this problem, the routine first calculates a denominator from the number of times the query substructure fragment can be found within itself. Then the number of

occurrences of the query fragment in the target structure was calculated. This result when divided by the above mentioned denominator gives a fragment occurrence frequency for the fragment in the target structure properly corrected for the any symmetry.

`aromatise()`

This routine calculates for each atom in a molecule whether it is adjacent to or contained within a six membered aromatic ring. If so the `aromatic` flag in the 'c' atom structure is labelled **true**. The routine does not pick up atoms which are adjacent to the class of aromatic 5-membered rings. Consequently this limited definition of the term *aromatic* is specific to this work.

`atom_count()`

This routine takes an atomic symbol (e.g. `Cl`) and searches for the number of occurrences of that symbol within a compound.

`gethazard()`

This is the key routine in the programs that predict a hazard index for compounds. The coefficients for each fragment are entered into the file containing this routine (`gethazard.c`). If one wishes to change the coefficients or even add in new type 2 fragment routines this is the file to edit.

Specialised Routines for Specific Fragment Subset

The type 2 fragment searching routines are shown in Table A.1. A computer coded representation of the molecular structure being studied is fed to the subroutine and it returns the number of occurrences

of the type 2 fragment in the molecule. The term `arom` implies that at least one of the atoms of the fragment either is, or is attached to, an atom that has its `aromatic` flag (set by the `aromatise()` routine) set to **true**. The number of aliphatic (i.e. non-aromatic) occurrences of a fragment is calculated from the difference between the total number of occurrences of the fragment and the number of occurrences of the fragment deemed aromatic.

Subroutine	Returns number of...
acrylate_count.c	acrylic ester fragments
aldehyde_count.c	aldehyde fragments
amide_count.c	amide fragments
amine_count.c	$1^o, 2^o$ or 3^o amine fragments
anhydride_count.c	anhydride fragments
anydouble_count.c	any double-bonded atom pairs
anytriple_count.c	any triple-bonded atom pairs
arom_COdouble_count.c	aromatic C=O fragments
arom_amine_count.c	aromatic $1^o, 2^o$ or 3^o amine fragments
arom_atom_count.c	aromatic occurrences of a specified atom
arom_carboxyl_count.c	aromatic carboxyl fragments
arom_nitro_count.c	aromatic nitro fragments
atom_count.c	occurrences of a specified atom
benzene_count.c	benzene ring structures
beta_lactam_count.c	β-lactam fragments
carboxyl_count.c	carboxyl fragments
CCdouble_count.c	C=C fragments
CNdouble_count.c	C=N fragments
CNtriple_count.c	C≡C fragments
COdouble_count.c	C=O fragments
ester_count.c	ester fragments
ethanolamine_count.c	ethanolamine fragments
ether_count.c	ether fragments
isocyanate_count.c	isocyanate fragments
ketone_count.c	ketone fragments
methyl_count.c	methyl fragments
nitro_count.c	nitro fragments
pyridine_count.c	pyridine fragments

Table A.1: Subroutines use to identify type 2 fragments in a compound.

Appendix B

CD-ROM

The accompanying CD-ROM has been created to contain both software and data described in this thesis. The CD-ROM uses an `iso9660` file-system and so should be accessible from the vast majority of operating systems. Please bare in mind the following notes:

1. The contents of the CD-ROM remain the intellectual property of the author and The University of Edinburgh. For further details read the file LICENSE.txt in the root directory.
2. For a full description of the contents of the CD-ROM read the file CONTENTS.txt
3. For general details relevant to each directory read the file README.txt in that directory.
4. Useful third-party software has been included in the `programs\other` directory. The conditions of use of the software should be read carefully before use.

Use of the CD-ROM is at the end-user's risk. Although every effort has been made to ensure the CD-ROM is virus free users should scan the disk before opening it.

The following data and programs are available on the CD-ROM:

1. `Hazassess` - the web based asthma hazard prediction program.
2. `batchpredict` - can predict hazards for multiple compounds (from an *.sdf file)
3. `cluster` - a program to generate clusters from a table of property data.
4. `qstruct` - a program to fragment all the compounds in one *.sdf file using type 1 fragments (stored in another fragment dictionary *.sdf file)
5. `fingerprint` - a program to fragment all the compounds in one *.sdf file using type 2 fragments
6. `kappa` - a small routine to calculate a kappa value given the four number entries to the 2x2 table.
7. Asthma and skin databases in proprietary format.
8. Asthma and skin databases in readable *.sdf text format.
9. Web based access to the database information.
10. Copy of the thesis including all figures in encapsulated postscript format.

Appendix C

Compendium of Chemicals

C.1 Asthmagens

Compounds listed in Table C.1 are those low molecular mass organic chemicals reported or suspected of causing asthma. Several are marked with question marks, indicating that the published evidence was not sufficient for the compound to be included in the analyses. However, in an attempt to be a definitive sensitive list of published asthmagens these suspect compounds have been included for completeness.

Table C.1: LMW Organic Asthmagens

CAS Number	Compound and Reference
514-10-3	Abietic acid / Colophony[172, 59]
64-19-7	⋆ Acetic acid [184]
79-10-7	⋆ Acrylic acid [108]
957-68-6	7-Amino-cephalosporanic acid [114]
111-41-1	Amino-ethyl ethanolamine [92, 90]

551-16-6	6-Amino penicillamic acid [113]
317-34-0	Aminophylline [133]
26787-78-0	Amoxicillin [115]
69-53-4	Ampicillin [113]
121-25-5	Amprolium hydrochloride [143]
123-77-3	Azodicarbonamide [138, 139]
8001-54-5	⋆ Benzalkonium chloride [82]
69-57-8	Benzyl penicillin [113]
	‡ 1,3-bis(cyanatomethyl)-cyclohexane prepolymer BIC
17095-24-8	Black GR Reactive Dye (BK-5) [64, 187]
2425-06-1	Captafol [110]
13466-78-9	3-Carene[137]
1390-65-4	⋆ Carmine [72, 73]
	‡ Ceftazidime
15686-71-2	Cephalexin [13]
127-65-1	Chloramine-T [83, 84, 85]
55-56-1	Chlorhexidine [86]
88-04-0	Chloroxylenol [81]
12238-08-3	Cibachrome Brilliant Scarlet 32 Reactive Dye [63]
51481-61-9	Cimetidine [121]
561-27-3	Diacetyl morphine [122]
62-73-7	Dichlorvos[112]
101-77-9	⋆ 4,4'-diamino-diphenyl-methane[101]
100-37-8	2-Diethyl-aminoethanol [93]
109-55-7	3-(Dimethylamino)-Propylamine [94]
108-01-0	Di-methyl ethanolamine [23]
117-81-7	Dioctyl phthalate [124]
	⋆ Direct Black [68]
49745-95-1	Dobutamine Chlorohydrate [188]
101-68-8	Diphenylmethane Di-isocyanate [155, 156, 153]
51811-44-0	Drimaren Brilliant Blue K-BL Reactive Dye [63]
13838-16-9	⋆ Enflurane [131]
141-43-5	Ethanolamine [91, 89]

7085-85-0	Ethyl cyanoacrylate [103]
	‡ Ethyl methacrylate
107-15-3	Ethylenediamine [95, 96, 91]
75-21-8	Ethylene oxide [144, 145]
55-38-9	Fenthion [112]
50-00-0	Formaldehyde[79, 78, 80]
98-00-0	Furfuryl alcohol [142]
111-30-8	Glutaraldehyde[75, 76]
	Glycyl compound[130]
70-30-4	Hexachlorophene [87]
14166-21-3	Hexahydrophthalic anhydride[48, 47, 52]
822-06-0	Hexamethylene di-isocyanate [30, 31]
100-97-0	Hexamethylene tetramine [91]
2746-19-2	Himic anhydride[49]
304-20-1	Hydralazine [126]
123-31-9	Hydroquinone [146]
25584-83-2	⋆ Hydroxy-propyl-acrylate [108]
	‡ Indigotine
	⋆ Isobutyl methacrylate [108]
54-85-3	Isoniazid[128]
78-59-1	⋆ Isophorone diamine[101]
4098-71-9	Isophorone di-isocyanate [32]
	⋆ Lanasol Yellow [67]
	⋆ Lauryl dimethyl benzyl ammonium chloride [81]
12226-52-7	Levafix Brilliant Yellow E36 Reactive Dye [63]
108-31-6	Maleic anhydride[50]
	⋆ Methacrylate [108]
148-01-6	2-Methyl-3,5-dinitro-benzamide [97]
555-30-6	α-Methyl DOPA[120]
137-05-3	Methyl cyanoacrylate [104, 103, 105, 108]
	⋆ Methyl hexahydrophthalic anhydride [51]
80-62-6	Methyl methacrylate [102, 103]
19438-64-3	⋆ Methyl tetrahydrophthalic anhydride [53, 54, 55]

7786-34-7	Mevinphos [111]
	MM22383 [123]
57-27-2	⋆ Morphine[122]
109-02-4	N-Methyl morpholine[99]
3173-72-6	1,5-Naphthalene di-isocyanate [34]
106-50-3	Para-phenylene diamine [100]
305-80-6	Pauli's reagent [140]
52-67-5	Penicillamine [129]
	‡ Penicillin G
39878-87-0	Phenyl glycine acid chloride [125]
85-44-9	Phthalic anhydride [41, 40, 42]
	‡ Piperacillin sodium
110-85-0	Piperazine [97, 91, 98] hardener
89-32-7	Pyromellitic dianhydride [40]
	Red-BBN Reactive Dye [64]
7531-92-2	Rosin / plicatic acid [189]
36519-31-0	Salbutamol [132]
123354-92-7	Sodium iso-nonanoyloxybenzenesulphonate [88, 116]
8025-81-8	Spiramycin [116, 117, 118]
100-42-5	Styrene [102, 134, 135, 136]
72-14-0	Sulphathiazole [127]
1897-45-6	Tetrachloro-isophthalonitrile [109]
117-08-8	Tetrachloro-phthalic anhydride [43, 44, 45, 40, 46]
60-54-8	Tetracycline [190, 119]
112-57-2	⋆ Tetraethylene-pentamine[101]
31330-63-9	1-(5-Tetrazoly)-4-guanyl tetrazene hydrate [141]
584-84-9	2,4-Toluene di-isocyanate [37, 28, 191, 38]
102-71-6	‡ Triethanolamine
	Trihydroxy-methyl-propyl-triacrylate [108]
	‡ Triglycidyl isocyanurate
552-30-7	Trimellitic anhydride [42, 47, 40, 48, 40, 49]
3779-63-3	1,3,5-tris-(6-isocyanato-hexyl)-<1,3,5>triazinane-2,4,6-trione [192, 193]

Key: ⋆ indicates possible asthmagen but rejected, omitted or missed from the original asthmagen dataset.

‡ indicates asthmagen in the validation set and derived from published literature from 1995 onwards.

† indicates asthmagen in the validation set but derived from the SWORD Scheme (see p57).

C.2 Skin Compounds

A number of the citations (particularly those in foreign language journals) in this appendix were consulted by title and abstract only. The chemical property data was obtained from various sources, principally the Beilstein Online Database [194], the Merck Index [195] and the Aldrich Catalogue [196]. Synonyms are delimited by a semi-colon.

1. Diphenyl Guanidine [197]
2. Ethyl alcohol; Ethanol [198]
3. Disulfiram; Tetraethylthiuram disulfide [199]
4. Propylene oxide [200]
5. Metanil yellow [201]
6. Cyanamide [202, 203]
7. Ethylene diamine; 1,2-diaminoethane [204, 205, 206, 207, 208, 209, 210, 211]
8. Diphenyl-thiourea [212, 213]
9. Spectinomycin [214]
10. Furazolidone [214]

11. Tylosin [215, 216, 217, 218]e
12. Triforine [219]
13. Trifluralin [220]
14. Benefin; Benfluralin [220]
15. Captan; 2-trichloromethylsulfanyl-3a,4,7,7a-tetrahydro-iso-indole-1,3-dione [221, 222, 223, 224, 225]
16. Benzydamine [226, 227]
17. 2-<N-ethyl-4-amino-3-methyl-anilino>-ethanol; CD4-colour developer [228]
18. Aziridine cross-linker cx100; Ethylene-imine; [229, 230]
19. CD-2 Colour developer; N',N-diethyl-2-methyl-p-phenylene diamine [228]
20. Propranolol [231, 232]
21. Thimerosal; Ethyl[2-mercaptobenzoato(2–)-O,S]-mercurate-(1–) sodium [233, 234, 225]
22. Triethylenetriamine; Trientine [209]
23. Limonene; 4-isopropenyl-1-methyl-cyclohexene; p-mentha-1,8-diene; [233, 235]
24. Cresyl glycidyl ether; 2-p-tolyloxymethyl-oxirane [206]
25. N-<1,3-dimethyl-butyl>-N'-phenyl-p-phenylenediamine; Santoflex 13 [236, 237]
26. Triglycidyl isocyanurate; 1,3,5-tris-oxiranylmethyl-<1,3,5>triazinane-2,4,6-trione [238, 239, 240, 241]
27. Diamino-diphenyl methane [242, 243, 155, 238, 153]s
28. 2-hydroxy-ethyl methacrylate; methacrylic acid-<2-hydroxy-ethyl ester>; 2-methyl-acrylic acid 2-hydroxy-ethyl ester [238, 244, 245, 246]

29. Benzene-1,4-diamine; p-Phenylenediamine [247, 248, 249, 250, 251, 234, 252]r

30. N-ethyl-4-toluene sulfonamide; N-ethyl-4-methyl-benzene-sulfonamide [253]

31. 4-tolyl diethanol-amine; N,N-bis-<2-hydroxy-ethyl>-p-toluidine [253]

32. Isophorone diamine; 3-amino-methyl-3,5,5-trimethyl-cyclohexylamine [254, 255, 256]

33. p-Toluenesulfonyl Chloride; Tosyl Chloride [257, 258]

34. Benzoyl peroxide[259, 260, 261]

35. Diethylenetriamine; N1-(2-amino-ethyl)-ethane-1,2-diamine [262, 209, 210, 263]

36. Bupirimate; dimethyl-sulfamic acid mono-(5-butyl-2-ethylamino-6-methyl-pyrimidin-4-yl) ester [264]

37. Benomyl [223]

38. 4-N,N-dimethyl-amino-benzene diazonium chloride [265]

39. Thiourea; Thiocarbamide [265, 266]

40. Methyl methacrylate [267, 268]

41. Falcarinol; heptadeca-1,9(Z)-diene-4,6-diyne-3-ol [269]

42. 2-methyl-isothiazol-3-one [270, 271, 272, 273]

43. Formaldehyde [270, 151, 152, 232]

44. Pyridine [274]

45. 2,2-Bis-(4-hydroxyphenyl)-propan-diacrylate [275]

46. Ethyl-2-cyanoacrylate [276]

47. Benzalkonium Chloride [277, 150, 278]

48. Diphenylmethane di-isocyanate; bis-(4-isocyanato-phenyl)-methane; 4,4'-methanediyl-bis-phenyl isocyanate [155, 156, 153]

49. Glutaraldehyde [150, 151, 152]

50. Kathon 930; 4,5-dichloro-2-octyl-isothiazol-3-one [279]

51. 3-(dimethylamino)-propylamine [280]

52. Alachlor; 2-chloro-N-(2,6-diethylphenyl)-N-methoxymethylacetamide [281]

53. Kathon 893; Octhilinone; 2-octyl-isothiazol-3-one [282, 273, 283]

54. N,N-diethyl-benzene-1,4-diamine [284]

55. N-Phenyl-N'-cyclohexylparaphenylenediamine; N-cyclohexyl-N'-phenyl-benzene-1,4-diamine [285]

56. Chloro-hydroxyimino-acetic acid ethyl ester; Oxalic acid-2-ethyl ester-1-chloride-1-oxime [286]

57. Benzo<d>isothiazol-3-one; [287, 288, 289, 273, 290, 291]

58. Nicergoline [292]

59. Cyclohexanone; Keto-hexamethylene [293]

60. Toluene di-isocyanate; 2,4-di-isocyanato toluene; 2,4-tolylene di-isocyanate; TDI; Nacconate 100; 4-methyl-1,3-phenylene di-isocyanate; [153]

61. Hexamethylene di-isocyanate [153, 154]

62. Ethyl-(ethoxymethylene)-cyanoacetate; 2-cyano-3-ethoxy-acrylic acid ethyl ester [294]

63. Triphenyl phosphite [295]

64. carvone; p-mentha-6,8-dien-2-one [296]

65. 2–Mercaptobenzothiazole; Mercaptobenzothiazole; Captax; [297, 298]

66. Dimethyl cyano-carbonimido-dithioate; N-cyano-dithio-carbonimidic acid dimethyl ester [299]

67. Minoxidil; 3-oxy-6-piperidin-1-yl-pyrimidine-2,4-diamine [300]

68. Dazomet; 3,5-dimethyl-<1,2,4>-thiadiazinane-2-thione [301]

69. 4-diazo-N,N-diethylaniline; 4-(diethylamino)benzenediazonium [302, 303]

70. Resorcinol; 1,3-benzenediol [248]

71. Pyrogallol; 1,2,3-Benzenetriol [248]

72. Nitroglycerin; 1,2,3-Propanetriol-trinitrate [304, 305]

73. Azaperone [306]

74. 1,1,1-Trichloroethane [307]

75. Tripropylene glycol diacrylate [308]

76. Trimethyl-oxiranylmethyl-ammonium chloride; Glycidyl trimethyl ammonium; [309, 310, 311]

77. Glycerol mono-thioglycolate [312, 313, 314]

78. 4-Chloro–7-nitro-benzofuran [315]

79. Mitomycin-C [316]

80. 2,2-Bis-(4-oxiranylmethoxy-phenyl)-propane; DGEBA; 2,2-bis-(4-glycidyloxyphenyl)-propane [317, 318]

81. C,C'-m-phenylene-bis-methylamine; m-xylylenediamine; 3-aminomethyl-benzylamine [319]

82. Methacrylic acid; Methacrylate; [244]

83. Acrylic acid; Acrylate; [244]

84. Kitasamycin [320]

85. Midecamycin (A1) [320]

86. Ethyl thiourea [213]

87. Indigo carmine [321]

88. Monensin [321]
89. Thiabendazole [321]
90. Amprolium hydrochloride [321]
91. Nitrofurazone [322, 323]
92. Olaquindox [324, 325]
93. Bisphenol A; Bisphenol A glycerolate (1 glycerol/pheno) diacrylate; [326, 327]
94. Styrene [328]
95. Butylated hydroxy-toluene [329]
96. Thiamin; Vitamin B1 [330]
97. 1,2-dichloro-propane [331]
98. 1,3-Dicyclohexyl-carbodiimide[332, 333]
99. Chloramine-T [334, 335, 152]
100. Dimethoate [223]
101. Carbaryl [223]
102. Thiram [223, 336]
103. Propineb [223]
104. Propachlor [223, 337]
105. Metham Sodium [223]
106. Tetrafluoroterephthalonitrile [338]
107. TFX diamine; 1,4-bis(aminomethyl)-2,3,5,6-tetrafluorobenzene; [338]
108. Tri-(ethylene glycol)-dimethacrylate [339]

109. 4-chloro-3,5-dimethyl-phenol; parachlorometaxylenol (PCMX); p-chloro-m,m'-xylenol [340, 341]

110. Tetracaine [342]

111. Dicyanodiamide [343]

112. Hexamethylenetetramine [344]

113. Oct-2-ynoic acid methyl ester [345]

114. Non-2-ynoic acid methyl ester [345]

115. *para*-Aminophenol[234]

116. *ortho*-Aminophenol [234]

117. Quinophthalone [234, 346]

118. Procaine hydrochloride; Novocaine [234, 232]

119. N,N-diethyl-beta-chloro-ethylamine [347]

120. Crystal Violet lactone [348]

121. Folpet [224]

122. Captafol; Difolatan; [224]

123. Ranitidine [349, 350]

124. N,N'-bis-<2,alpha-dichloro-benzylidene>-hydrazine [351]

125. diazonium compound [352]

126. 2-hydroxyethyl-acrylate [353]

127. Trimethylolpropane triacrylate [354, 355, 356, 357, 358]

128. Diglycidylhexahydrophthalate; Cyclohexane-1,2-dicarboxylic acid dioxiranylmethyl ester [157]

129. Phenyl salicylate [359]

130. Aminotriazole [360]

131. Ortho-benzyl-parachlorophenol; Clorophene [361]

132. 5-chloro-2-methyl-isothiazol-3-one; Kathon CG; [272, 273, 225]

133. (1-methylethylidene) bis <4,1-phenyleneoxy(2-hydroxy-3,1-propanediyl)> bismethacrylate; 2-methyl-acrylic acid 2-hydroxy-3-<4-(1-(4-<2-hydroxy-3-(2-methyl-acryloyloxy) -propoxy>-phenyl)-1-methyl-ethyl)-phenoxy>-propyl ester [362]

134. Dodecyl-di-(aminoethyl)-glycine; Dodicin; Tego 51; Ampholyt-G [362, 154, 363, 364, 365, 366]

135. Chemical Mace; omega-chloroacetophenone; alpha-chloroacetophenone [367, 368, 369]

136. Aminoethyl-ethanolamine [370]

137. Neomycin [371, 372]

138. Streptomycin [372, 373]

139. 1-Naphthyl iso-thiocyanate [374]

140. 2-{5-dimethyl-aminomethyl-furan-2-ylmethylsulfanyl)-ethylamine [375, 350]

141. Mercaptoacetic acid [376]

142. Tetra-(ethylene glycol)-diacrylate [377, 378]

143. 1,3-dibutylthiourea [379]

144. Tetryl; Nitramine; N-methyl-2,4,6,N-tetranitro-aniline [380]

145. TNT; 2,4,6-Trinitro-toluene [380]

146. bis-acryloylamino-methane; [245, 381, 382]

147. Epichlorhydrin [383]

148. Methylene bis(4-cyclohexylisocyanate) [384, 385]

149. Tetrachloroisophthalonitrile [386]

150. Pentaerythritol triacrylate [356, 357, 387, 358]

151. 1,6-Hexanediol diacrylate [356, 358]

152. 4-tert-Butyl-catechol [388]

153. 5-amino-4-chloro-2-phenyl-2H-pyridazin-3-one; chloridazon; pyrazon; [389]

154. Sorbitan Monooleate; Span(R); Tween(R) [390]

155. 2-hydroxymethyl-2-nitro-propane-1,3-diol [385]

156. 2-bromo-2-nitro-propane-1,3-diol; bronopol [385]

157. Methyl-dithio-carbamic acid [391]

158. 1,3-diethyl-thiourea; N,N'-diethyl-thiourea [392]

159. 4,7-dichloro-quinoline [393]

160. Bis-acryloylaminomethyl-ether; N,N'-[oxybis(methylene)]bis-2-propenamide [381]

161. N-<2-hydroxy-ethyl>-N-methyl-acrylamide; N-[(2-hydroxyethoxy)-methyl]-2-propenamide [381]

162. (2-hydroxyethyl)dimethylsulphoxonium ion [394]

163. Benzidine [395]

164. 4-bromomethyl-6,8-dimethyl-2(1H)-quinolone [396]

165. 4-bromoacetoacet-2,4-dimethylanilide; omega-bromo-2,4-acetoacetxylidide; 4-bromo-N-(2,4-dimethyl-phenyl)-3-oxo-butyramide [396]

166. 3-<(4-methyl-piperazin-1-ylimino)-methyl>-rifamycin; Rifampicin [397]

167. diphenyl-cyclopropenone; 2,3-diphenyl-cycloprop-2-enone [398]

168. Dimethoxane; 4-acetoxy-2,6-dimethyl-<1,3>dioxane; Acetic acid 2,6-dimethyl-<1,3>dioxan-4-yl ester [399]

169. (2-chloro-acetylamino)-methanol; 2-chloro-N-hydroxymethyl-acetamide; chloro-acetic acid-<hydroxymethyl-amide> [400]

170. R-3,4-Dimethoxydalbergione; 2,3-dimethoxy-5-(1-phenyl-allyl)-<1,4>benzoquinone; [401]

171. Spiramycin 1[218, 117]

172. Penethamate BP [218]

173. Methyl isothiocyanate [402]

174. Eugenol [403, 232]

175. Aminophylline [232]

176. Chlorpromazine hydrochloride [232]

177. N-methyl-para-aminophenol [404]

178. 2-hydroxy-3-naphtoic acid-o-aniside [404]

179. N-acetoacetyl-benzylamide [404]

180. Monosulfiram [405]

181. N-isopropyl-N'-phenyl-p-phenylenediamine [406, 237, 407]

182. Ethyl acrylate [408]

183. Helenalin; 6-alpha,8-beta-dihydroxy-4-oxo-ambrosa-2,11(13)-dien-12-oic acid-8-lactone [409]

184. Carabron; 6a-methyl-3-methylene-1-(3-oxo-butyl)-octahydro-5-oxa-cyclopropa<f>inden-4-one [409]

185. Ampicillin; Sodium <2S-<2-alpha,5-alpha,6-beta(S*)>>-6-(aminophenylacetamido) -3,3-dimethyl-7-oxo-4-thia-1-azabicyclo<3.2.0>heptane-2-carboxylate [410]

186. Oxacillin; 6-beta-(5-methyl-3-phenyl-isoxazole-4-carbonylamino)-penicillanic acid; (5-methyl-3-phenyl-isoxazol-4-yl)-penicillin [410]

187. N,N-dimethyl-para-phenylenediamine; N,N-dimethyl-benzene-1,4-diamine; N,N-dimethyl-benzene-1,4-diamine [411]

188. Apomorphine [412]

189. 4-tert-butyl-phenol; Para-tert-butyl-phenol. [413]

190. Sodium amidotrizoate; Diatrizoate sodium [414]

191. Pentaerythritol tetrakis-3-mercaptopropionate [415]

192. 3-mercaptopropionic acid [415]

193. 2,2',2"-<1,3,5>triazinane-1,3,5-triyl-tris-ethanol; 1,3,5-tris-<2-hydroxy-ethyl>-hexahydro-<1,3,5>triazine; 2-<3,5-bis-(2-hydroxy-ethyl)-<1,3,5>triazinan-1-yl>-ethanol; [416]

194. Chloracetamide [417]

195. Nonylphenol ethoxylate; nonoxynol-6 [418]

196. Squaric acid dibutyl ester [419]

197. Naphthol AS; 3-hydroxy-naphthalene-2-carboxylic acid anilide; 3-hydroxy-<2>naphthoic acid anilide; 3-hydroxy-naphthalene-2-carboxylic acid phenylamide [420]

198. 4-vinyl-1-cyclohexene diepoxide; 1,2-epoxy-4-oxiranyl-cyclohexane; 3-oxiranyl-7-oxa-bicyclo<4.1.0>heptane [421]

199. Bromo-<4-methoxy-phenyl>-acetic acid ethyl ester [422]

200. 3-dodecyloxy-propylamine; laurixamine; 3-dodecyloxy-propylamine [423]

C.3 Controls

1,1,1-Tetrachloro-2,2-difluoroethane
1,1,1-Trichloroethane
1,1,1-Trichlorobis(chlorophenyl)ethane
1,1,2,2-Tetrabromoethane

1,1,2,2-Tetrachloro-1,2-difluoroethane
1,1,2-Trichlorotrifluoroethane
1,1-dichloroethane
1,2,3-Trichloropropane
1,2,3-Trimethylbenzene
1,2,4-Trichlorobenzene
1,2,4-Trimethylbenzene
1,2-Dibromoethane
1,2-Dichloroethylene
1,2-Dinitrobenzene 1,2-dichlorobenzene
1,3,5-Trimethylbenzene
1,3-Dichloro-5,5-dimethyl-hydantoin
1,3-Dimethylbutyl acetate
1,3-dinitrobenzene
1,4-Dichlorobenzene
1,4-Dinitrobenzene
1,4-Dioxane
1,6-Hexanolactam
1-Methoxypropan-2-ol
1-Methyl-2-pyrrolidone
1-Methylbutyl acetate
1-Nitropropane
2,2'-Dichloro-4,4'methylene dianiline
2,2'-Iminodi(ethylamine)
2,2'-Iminodiethanol
2,2'-Oxydiethanol
2,3-Epoxypropyl isopropyl ether
2,3-Xylidene
2,4,5-T 2,4,6-Trinitrotoluene
2,4-D (ISO) 2,4-Xylidine
2,5-Xylidene
2,6-Dimethyl heptan-4-one
2,6-Ditertiary-butyl-para-cresol
2,6-Xylidene
2-Aminoethanol
2-Butoxyethanol
2-Chloro-6-(trichloromethyl)pyridine
2-Chlorobuta-1,3-diene
2-Clorotoluene
2-Ethoxy-ethyl acetate
2-Ethoxyethanol
2-Ethylhexyl chloroformate
2-Furaldehyde
2-Hydroxy propyl acrylate
2-Methoxyethylacetate
2-Methyl-4,6-dinitro phenol
2-Methylcyclohexanol
2-Methylcyclohexanone
2-Methylpentane-2,4-diol
2-Methylpropan-1-ol
2-Methylpropan-2-ol
2-Nitro toluene
2-Phenylpropene
2-Pyridylamine
2-chloroethanol
2-sec-Butyl phenol
3,4-Xylidene
3,5,5-Trimethylcyclohex-2-enone
3,5-Xylidene
3-Methyl cyclo hexanol
3-Methylbutan-1-ol
3-Methylstyrene
3-Nitro toluene
4-Hydroxy-4-methyl-pentan-2-one
4-Methyl cyclo hexanol
4-Methylpent-3-ene-2-one

4-Methylpentan-2-ol
4-Methylpentan-2-one
4-Methylstyrene
4-Nitro toluene 4-Nitroaniline 4-Ethyl-morpholine 5-
5-Methylhexan-2-one
6,6-Di-tert-butyl-4,4-thiodi-meta-cresol
Acetic anhydride
Acetone
Acetonitrile
Acrylaldehyde
Acrylamide
Acrylic acid
Acrylonitrile
Aldrin
Allyl alcohol
Allyl-2,3-epoxypropyl ether
Azinphos-methyl
Benomyl
Benzene
Benzenethiol
Benzyl butyl phthalate
Biphenyl
Bis(2,3-epoxypropyl) ether
Bis-(chloromethyl)-ether
Bornan-2-one
Bromacil
Bromochloromethane
Bromoethane
Bromoform
Bromomethane
Bromotrifluoromethane
Buta-1,3-diene
Butan-1-ol
Butan-2-ol
Butan-2-one
Butyl acetate
Butyl acrylate
Butyl lactate
Captan
Carbaryl
Carbofuran
Carbon tetrabromide
Carbon tetrachloride
Chloroacetophenone
Chloroactealdehyde
Chlorobenzene
Chlorodifluoromethane
Chloroethane
Chloroform
Chloromethane
Chloropentafluoroethane
Chloropyrifos
Cryofluorane
Cumene
Cyanamide
Cyanogen Chloride
Cyclohexane
Cyclohexanol
Cyclohexanone
Cyclohexene
Cyclohexylamine
Di-n-butyl hydrogen phosphate
Diallyl phthalate
Diazinon
Dibenzoyl peroxide
Dibromodifluoromethane

Dibutyl phthalate
Dichlorodifluoromethane
Dichloromethane
Dicyclohexylphthalate
Dieldrin
Diethyl phthalate
Di-isobutylphthalate
Di-isopropylamine
Dimethyl formamide
Dimethylamine
Diphenyl amine
Diquat dibromide
Diuron
Ethane thiol
Ethanol
Ethyl acrylate
Ethyl chloro formate
Ethylamine
Fenchlorphos
Formic acid
Glycerol trinitrate
Heptan-3-one
Hexahydro-1,3,5-trinitro-1,3,5-triazine
Indene
Isobutyl acetate
Isopropyl acetate
Ketene
Mequinol
Methacrylic acid
Methane thiol
Methomyl
Methoxyethanol
Methyl acrylate
Dichloroacetylene
Dichlorofluoromethane
Dichlorvos
Dicyclopentadiene
Diethyl ether
Diethylamine
Di-isopropyl ether
Dimethoxymethane
Dimethyl phthalate
Dioxathion
Diphenyl ether
Disulfoton
Endosulfan
Ethane-1,2-diol
Ethyl acetate
Ethyl benzene
Ethyl formate
Ethylene dinitrate
Formamide
Glycerol
Heptan-2-one
Hexachloroethane
Hydrogen cyanide
Iodoform
Isopentyl acetate
Isopropyl chloroformate
Malathion
Mercaptoacetic acid
Methacrylonitrile
Methanol
Methoxychlor
Methyl acetate
Methyl formate

Methylamine
Mevinphos
Morpholine
N-Methyl-N,2,4,6-tetra-nitroaniline
NN-Dimethylaniline
Naled
Nicotine
Nitrobenzene
Octachloronaphthalene
Oxalonitrile
Parathion
Pentachlorophenol
Pentan-2-one
Pentane
Phenol
Phorate
Picloram
Piperidine
Propan-1-ol
Propane-1,2-diol
Propoxur
Pyridine
Resorcinol
Sodium fluoro-acetate
Strychnine
Sulfotep
Tetrachloro ethylene
Tetrahydrofuran
Tetramethyl succinonitrile
Toluene
Tri-(ortho-Tolyl) phosphate
Trichlorofluoromethane
Triethylamine
Methylcyclohexane
Monochloroacetic acid
N-Methyl aniline
NN-Dimethyl acetamide
NN-Dimethylethylamine
Naphthalene
Nitro-ethane
Nitromethane
Oxalic acid
Paraquat dichloride
Parathion-methyl
Pentaerythritol
Pentan-3-one
Pentyl acetate
Phenyl-2,3-epoxypropyl ether
Phosgene
Picric acid
Prop-2-yn-1-ol
Propan-2-ol
Propionic acid
Propylene-1,2-dinitrate
Pyrocatechol
Rotenone
Sodium-2-(2,4-dichlorophenoxy)ethyl sulphate
Sucrose
TEPP
Tetraethyl orthosilicate
Tetramethyl orthosilicate
Thiram
Tri-(n-butyl) phosphate
Trichloroethylene
Trichloronitromethane
Trimethyl phosphite

Trimethylamine
Vinyl acetate
Vinylidene chloride
gamma-BHC
meta-Cresol
n-Butyl chloroformate
n-Butylamine
n-Hexane
n-Propyl acetate
ortho-Xylene
ortho-Cresol
para-Anisidine
para-Toluenesulphonyl chloride
tert-Butyl acetate
Triphenyl phosphate
Vinyl chloride
Warfarin
meta-Xylene
n-Butane
n-Butyl glycidyl ether
n-Heptane
n-Octane
ortho-Acetyl-salicylic acid
ortho-Anisidine
para-Xylene
para-Cresol
sec-Butyl acetate
Hexan-2-one

Appendix D

Miscellaneous

D.1 Miscellaneous Tables

Compound	Validation Category	Reason for Exclusion
Iodomethane	Control	Control in original dataset
Cumene	Control	Control in original dataset
Chlorobenzene	Control	Control in original dataset
2-Ethoxyethanol	Control	Control in original dataset
2-Pyridyl amine	Control	Control in original dataset
Hexan-2-one	Control	Control in original dataset
N,N-Dimethethylamine	Control	Control in original dataset
Methomyl	Control	Control in original dataset
Para-Benzoquinone	Control	Asthmagen in original dataset
Dichlorvos	SWORD	Asthmagen in original dataset
Acetic acid	SWORD	Asthmagen in original dataset
Chlorothalonil	SWORD	Asthmagen in original dataset
Triethylene tetramine	SWORD	Asthmagen in original dataset

Table D.1: Compounds rejected from the validation.
A number of compounds were rejected from the validation dataset provided by Dr Agius. The rejection of validation compounds might be considered a source of experimenter bias so the justification for their exclusion is listed here.

D.2 Interpretation of Logistic Regression Tables

The initial interpretation of Tables 3.14 and 3.16 may lead one to the conclusion that certain variables should not be included because the standard error of β is of a larger magnitude than the variable's coefficient β. The interpretation of the standard error should be that it gives a measure of the possible range of values the coefficient β could take but should not be the determinant of whether the variable should be included in the model. Indeed, when the independent variable predicts too well then the problem of complete separation occurs [186]. This leads to very large coefficients and standard errors. The inclusion of a variable in the model is based upon a statistically significant reduction in the ability of the model to predict accurately. This is implicit in the use of the backward log likelihood statistic as the determinant of

Compound Name	Hazard	Reason for exclusion from analysis
Acetic acid (see Figure	0.83	This compound was missed. Entered and used as a control
		and in the later analyses but a report of asthma was found [184]
Acrylic acid	0.15	This compound is a suspected asthmagen [108] which was used as a control in the analysis.
Methyl-tetrahydrophthalic anhydride	0.93	Mixture of three isomers. All have the same calculated asthma hazard.
Methyl-hexahydrophthalic anhydride	0.93	Not a full paper.
Tetraethylene pentamine	1.00	Constituent of EPO 60, Lambourn *et al.* described occupational asthma caused by EPO 60 [101]
Isophorone diamine	0.81	Another constituent of EPO 60 [101].
4-4' diamino-diphenyl methane	0.22	Another constituent of EPO 60 [101].
Enflurane	0.00	Paper missed due to lack of 'occupational' or equivalent keyword [131].

Table D.2: Possible asthmagens excluded from the study.
Several potential asthmagens were excluded from the study. As this might be considered a source of experimenter bias they are documented here complete with the reason for their exclusion.

whether variables are included in the model.

D.3 Abbreviations

C.I.'s	Confidence Intervals
HDI	Hexamethylene Di-isocyanate
HOR	Hazard Odds Ratio (see page 73)
IgE	Immunoglobulin E
IgG	Immunoglobulin G
LMW	Low Molecular Weight
MDL	Molecular Design Limited (see `http://www.mdli.co.uk/`)
MIDAS	Manchester Information Datasets and Associated Services (see `http://ww`
MTHPA	Methyl Tetrahydro Phthalic Anhydride
OA	Occupational Asthma
SAR	Structure-Activity Relationship
TDI	Toluene Di-isocyanate

References

Note: References cited in the main text have been properly consulted. References cited in the appendices may have only had the title and abstract reviewed in order to confirm they refer to cases of occupational sensitisation. This is particularly the case where several reports refer to a single chemical. They are included here in order to provide a comprehensive listing for the chemical case data.

[1] C. H Hennekens and J. E. Buring. *Epidemiology in Medicine*. Little, Brown And Company., Boston/Toronto, first edition, 1987.

[2] G. M Downs and P. Willett. *The Use of Similarity and Clustering Techniques for the Prediction of Molecular Property Data*, pages 247–279. ECSC, EEC, EAEC, Brussels, 1991.

[3] C. K. Mathews and K. E. van Holde. *Biochemistry*, chapter 14, page 474. Benjamin/Cummings Publishing Company, Inc., 1990.

[4] B. W. Richardson. *Snow on Cholera*. Hafner, New York, 1965.

[5] A. Bradford Hill. *Principles of Medical Statistics*. The Lancet Ltd, London, 1971.

[6] D. Hunter. *The Diseases of Occupations*. The English Universities Press, London, 1975.

[7] M. R. Becklake. *Asthma in the Workplace*, chapter Epidemiology: Prevalence and Determinants, page 30. Dekker, 1993.

[8] B. T. Butcher and J. E. Salvaggio. Occupational asthma. *The Journal of Allergy and Clinical Immunology*, 78:547–556, 1986.

[9] M. Makins, editor. *Collins Softback English Dictionary*. Harper Collins, Aylesbury, England, 1993.

[10] I. Roit, J. Brostoff, and D Male. *Immunology*. Mosby, London, fourth edition edition, 1996.

[11] D. J. Hendrick. Occupational asthma - problems of definition. *Journal of Occupational Medicine*, 25(6):488–489, 1983.

[12] D. J. Hendrick. Address at Occupational Asthma - Current Trends in Prediction, Regulation and Clinical Assessment, The Members' House, Edinburgh Zoo. March 12 1998.

[13] I. Coutts, M. Dally, A. J. Newman-Taylor, C. Pickering, and N. Horsfield. Asthma in workers manufacturing cephalosporins. *British Medical Journal*, 283:950, 1981.

[14] J. T. Hodgson, J. R. Jones, R. C. Elliott, and J. Osman. *Self-reported work-related illness*. Health & Safety Executive., 1994.

[15] C. M. Fraser, J. A. Bergeron, A Mays, and S. E. Aiello, editors. *Merck Veterinary Manual*, chapter Immune System. Rahway, N.J., U.S.A., 1992.

[16] X-D. Zhang. *Relationship between chemical structure and airway sensitizing potential for organic acid anhydrides. An animal model*. PhD thesis, Lund, Sweden, 1997.

[17] I. Kimber. Assassment of allergic potential. *Drug Information Journal*, 30:287–292, 1996.

[18] F. Deschamps, A. Prevost, F. Lavaud, and S. Kochman. Mechanisms of occupational asthma induced by isocyanates. *Annals of Occupational Hygiene*, 42(1):33–36, 1998.

[19] R. J. Dearman, D. A. Basketter, and I. Kimber. Differential cytokine production following chronic exposure of mice to chemical respiratory and contact allergens. *Immunology*, 86:545–550, 1995.

[20] I. Kimber, J. Hilton, D. A. Basketter, and R. J. Dearman. Predictive testing for respiratory sensitization in the mouse. *Toxicology Letters*, 86:193–198, 1996.

[21] J. E. Fish. Occupational asthma: a spectrum of acute respiratory disorders. *Journal of Occupational Medicine*, 24:379–386, 1982.

[22] J. Nielsen, H. Welinder, A. Schutz, and S. Skerfving. *J. Allergy Clin. Immnol.*, 82:126–133, 1988.

[23] M. Vallieres, D. Cockcroft, D. Taylor, J. Dolovich, and F. Hargreave. Dimethyl ethanolamine-induced asthma. *American review of Respiratory Disease*, 115:867–871, 1977.

[24] H. P. Rang and M. M. Dale. *Pharmacology*, chapter 17. Churchill Livingstone, 1991.

[25] A Cartier and J. Malo. *Asthma in the Workplace*, chapter Occupational Challenge Tests, pages 234–237. Dekker, 1993.

[26] S. R. Durham, B. J. Graneek, R. Hawkins, and A. J. Newman Taylor. The temporal relationships between the increases in airway responsiveness to histamine and late asthmatic responses induced by occupational agents. *Journal of Allergy and Clinical Immunology*, 79:398–406, 1987.

[27] D. J. Ross, B. A. Sallie, and J. C. McDonald. SWORD'95: surveillance of work-related and occupational respiratory disease in the UK. *Occupational Medicine*, 45:175–178, 1995.

[28] L. M. Fabbri, D. Danieli, S. Crescioli, P. Bevilacqua, S. Meli, M. Saetta, and C. Mapp. Fatal asthma in a subject sensitized to toluene diisocyanate. *Am.Rev.Resp.Dis.*, 137:1494–1498, 1988.

[29] *Kirk-Othmer Encyclopedia of Chemical Toxicology*, volume 12, chapter Isocyanates, pages 45–64. John Wiley & Sons, Inc., New York, London, Sydney.

[30] J. L. Malo, G. Ouimet, A. Cartier, D. Levitz, and C. R. Zeiss. Combined alveolitis and asthma due to hexamethylene diisocyanate (HDI), with demonstration of crossed respiratory and immunologic reactivities to di-phenylmethane diisocyanate (MDI). *J Allergy Clin Immunol*, 72:413–419, 1983.

[31] L. Grammer, P. Eggum, M. Silverstein, M. Shaughnessy, J. Liotta, and R. Patterson. Prospective immunologic and clinical study of a population exposed to hexamethylene diisocyanate. *J Allergy Clin Immunol*, 82:627–633, 1988.

[32] C. Clarke and P. Aldons. Isophorone diisocyanate induced respiratory disease (IPDI). *Aust N Z J Med*, 11:290–292, 1981.

[33] M. Zammit-Tabona, M. Sherkin, K. Kijek, H. Chan, and M. Chan-Yeung. Asthma caused by diphenylmethane diisocyanate in foundry workers. *Am. Rev. Respir. Dis.*, 128:226–230, 1983.

[34] M. G. Harries, P. Burge Sherwood, M. Samson, A. J. Newman-Taylor, and J. Pepys. Isocyanate asthma: respiratory symptoms due to 1,5 naphthylene diisocyanate. *Thorax*, 34:762–766, 1979.

[35] NIH Technology Assessment Workshop Panel. The persian gulf experiance and health. *Journal of the American Medical Association*, 272:391–396, 1998.

[36] H. Savolainen. Antitrypsin phenotypes in occupational organic diisocyanate asthma. *Research Communication in Chemical Pathology and Pharmacology*, 71(3):385–386, 1991.

[37] D. Banks, R. Rando, and H. Barkman. Persistence of toluene diisocyanate-induced asthma despite negligible workplace exposure. *Chest*, 97:121–125, 1990.

[38] G. Pisati, A. Baruffini, and S. Zedda. Toluene diisocyanate induced asthma: outcome according to persistence or cessation of exposure. *Br J Ind Med*, 50:60–64, 1993.

[39] P. L. Paggiaro, A. M. Loi, O. Rossi, B. Ferrante, F. Pardi, M. G. Roselli, and L. Baschieri. Follow-up study of patients with respiratory disease due to toluene diisocyanate (TDI). *Clinical Allergy*, 14:463–469, 1984.

[40] K. M. Venables. Low molecular weight chemicals, hypersensitivity, direct toxicity: The acid anhydrides. *British Journal of Industrial Medicine*, 46:222–232, 1989.

[41] M. Wernfors, J. Nielsen, A. Schutz, and S Skerfving. Phthalic anhydride-induced occupational asthma. *Int. Archs. Allergy appl. Immun.*, 79:77–82, 1986.

[42] D. I. Bernstein, R. Patterson, and C. R. Zeiss. Clinical and immunological evaluation of trimellitic anhydride- and phthalic anhydride-exposed workers using a questionnaire with comparative analysis of enzyme-linked immunosorbent and radioimmunoassay studies. *Journal of Allergy and Clinical immunology*, 69(3):311–318, 1982.

[43] W. Howe, K. M. Venables, M. D. Topping, M. B. Dally, R. Hawkins, J. S. Law, and A. Newman Taylor. Tetrachlorophthalic anhydride asthma: evidence for specific igE antibody. *Journal of Allergy and Clinical Immunology*, 71(1):5–11, 1983.

[44] K. M. Venables, M. Topping, W. Howe, C. Luczynska, R. Hawkins, and A. J. Newman-Taylor. Interaction of smoking and atopy in producing specific igE antibody against a hapten protein conjugate. *British Medical Journal*, 290:201–204, 1985.

[45] K. M. Venables, M. Topping, A. J. Nunn, W. Howe, and A. J. Newman-Taylor. Immunologic and functional consequences of chemical (tetrachlorophthalic anhydride)-induced asthma after four years of avoidance of exposure. *Journal of Allergy and Clinical Immunology*, 80(2):212–218, 1987.

[46] K. M. Venables and A. J. Newman Taylor. Exposure-response relationships in asthma caused by tetrachlorophthalic anhydride. *Journal of Allergy and Clinical Immunology*, 85(1):55–59, 1990.

[47] D. R Moller, J. S. Gallagher, D. I. Bernstein, Wilcox T. G., H. E. Burroughs, and I. L. Bernstein. Detection of igE-mediated respiratory sensitization in workers exposed to hexhydrophthalic anhydride. *Journal of Allergy and Clinical Immunology*, 75(6):663–672, 1985.

[48] C. B. E. Chee, H. S. Lee, T. H. Cheong, Y. T Wang, and S. C. Poh. Occupational asthma due to hexahydrophthalic anhydride: a case report. *British Journal of Industrial Medicine*, 48:643–645, 1991.

[49] K. D. Rosenman, D. I. Bernstein, K. O'Leary, J. S. Gallagher, and L. D'Souza. Occupational asthma caused by himic anhydride. *Scandinavian Journal of Work, Environment & Health*, 13:150–154, 1987.

[50] H. S. Lee, Y. T. Wang, T. H. Cheong, K. T. Tan, B. E. Chee, and K. Narendran. Occupational asthma due to maleic anhydride. *British Journal of Industrial Medicine*, 48:283–285, 1991.

[51] H Keskinen, H. Nordman, O. Tupasela, E. Vaheri, P. Pfaffli, and M. Sarjanen. Methylhexahydrophthalic anhydride(MHHPA) induced asthma and rhinitis. *New England Reg. Allergy. Proc.*, 9:397, 1988.

[52] L. C. Grammer, M. A. Shaughnessy, and M. Lowenthal. Hemorrhagic rhinitis. an immunologic disease due to hexahydrophthalic anhydride. *Chest*, 104(6):1792–1794, 1993.

[53] J. Welinder H. Nielsen, C. Gustavsson, and I. Bensryd. Specific antibodies to methyl tetrahydrophthalic anhydride in exposed workers. *Clinical and Experimental Allergy*, 20:639–645, 1990.

[54] J. Nielsen, H. Welinder, and S. Skerfving. Allergic airway disease caused by methyl tetrahydrophthalic anhydride in epoxy resin. *Scandinavian Journal of Work, Environment & Health*, 15:154–155, 1989.

[55] J. Nielsen, H. Welinder, V. Horstmann, and S. Skerfving. Allergy to methyltetrahydrophthalic anhydride in epoxy resin workers. *British Journal of Industrial Medicine*, 49:769–775, 1992.

[56] R. J. Dearman and I. Kimber. Divergent immune responses to respiratory and contact chemical allergens: antibody elicited by phthalic anhydride and oxazolone. *Clinical and Experimental Allergy*, 22:241–250, 1992.

[57] X. Welinder H. Zhang, C. Gustavsson, B Bjork, and S. Skerfving. Structure-activity relationships of organic acid anhydrides as antigens in an animal model. *Toxicology*, 103:127–136, 1995.

[58] S. Sadhra, I. S. Foulds, C. N. Gray, D. Koh, and K. Gardiner. Colophony - uses, health effects, airborne measurements and analysis. *Annals of Occupational Hygiene*, 38:385–396, 1994.

[59] P. S. Burge, A. Wieland, A. S. Robertson, and D. Weir. Occupational asthma due to unheated colophony. *British Journal of Industrial Medicine*, 43:559–560, 1986.

[60] J. L. Malo, A. Cartier, J. L'Archeveque, C. Trudea, J. P. Courteau, and L. Bherer. Prevalence of occupational asthma among workers exposed to eastern white cedar. *American Journal of Respiratory & Critical Care Medicine*, 150:1697–1701, 1994.

[61] M. Chan-Yeung, S. Lam, and S. Koener. Clinical features and natural history of occupational asthma due to western red cedar (thuja plicata). *American Journal of Medicine*, 72:411–415, 1982.

[62] M. Chan-Yeung. Mechanism of occupational asthma due to western red cedar. *American Journal of Industrial Medicine*, 25:13–18, 1994.

[63] K. Alanko, H. Keskinen, F. Bjorksten, and S. Ojanen. Immediate-type hypersensitivity to reactive dyes. *Clinical Allergy*, 8:25–31, 1978.

[64] H. S. Park, Y. J. Kim, M. K. Lee, and C. S. Hong. Occupational asthma and igE antibodies to reactive dyes. *Yonsei Medical Journal*, 30:298–304, 1989.

[65] H. Park, M. Lee, B. Kim, K. Lee, J. Roh, Y. Moon, and C. Hong. Clinical and immunologic evaluations of reactive dye-exposed workers. *J Allergy Clin Immunol*, 87:639–649, 1991.

[66] C. S. Hong and H. S. Park. Heterogeneity of igE antibody response to reactive dye in sera from four different sensitised workers. *Clinical And Experimental Allergy*, 22:606–610, 1992.

[67] A new case of occupational asthma from reactive dyes with severe anaphylactic response to the specific challenge. *Am.J.Ind.Med.*, 21:209–216, 1992.

[68] A. Noferi. Su alcuni casi di asma bronchiale da nero diretto. *Folia Allergologica*, 14(6):534–537, 1967.

[69] A. Docker, J. M. Wattie, M. D. Topping, C. M. Luczynska, A. J. Newman Taylor, C. A. C. Pickering, P. Thomas, and D. Gompertz. Clinical and immunological investigations of respiratory disease in workers using reactive dyes. *British Journal of Industrial Medicine.*, 44:534–541, 1987.

[70] H. Keskinen, H. Nordman, and E. O. Terho. ECG ink as a cause of asthma. *Allergy*, 36:275–276, 1981.

[71] D. Rodenstein and D. Stanescu. Bronchial asthma following exposure to ECG ink. *Annals of Allergy*, 48:351–352, 1982.

[72] S. Quirce, M. Cuevas, J. M. Olaguibel, and A. I. Tabar. Occupational asthma and immunologic responses induced by inhaled carmine among employees at a factory making natural dyes. *J Allergy Clin Immunol*, 93:44–52, 1994.

[73] P. S. Burge, I. M. O'Brien, M. G. Harries, and J. Pepys. Occupational asthma due to inhaled carmine. *Clin.Allergy*, 9:185–189, 1979.

[74] P. M. B. Walker, editor. *Chamber Science and Technology Dictionary*. Chambers, Edinburgh, 1988.

[75] O. J. Corrado, J. Osman, and R. J. Davies. Asthma and rhinitis after exposure to glutaraldehyde in endoscopy units. *Human Toxicology*, 5:325–327, 1986.

[76] S. Jachuck, P. Bound, J. Steel, and P. Blain. Occupational hazard in hospital staff exposed to 2 per cent glutaraldehyde in an endoscopy unit. *J.Soc.Occup.Med.*, 39:69–71, 1989.

[77] J. Hilton, R. J. Dearman, D. A. Basketter, E. W. Scholes, and I. Kimber. Experimental assessment of the sensitising properties of formaldehyde. *Food and Chemical Toxicology*, 34:571–578, 1996.

[78] G. Uba, D. Pachorek, J. Bernstein, D. Garabrant, J. Balmes, W. Wright, and R. Amar. Prospective study of respiratory effects of formaldehyde among healthy and asthmatic medical students. *American Journal of Industrial Medicine*, 15:91–101, 1989.

[79] D. J. Hendrick and D. J. Lane. Formalin asthma in hospital staff. *British Medical Journal*, 1:607–608, 1975.

[80] P. S. Burge, M. G. Harries, W. K. Lam, I. M. O'Brien, and P. A. Patchett. Occupational asthma due to formaldehyde. *Thorax*, 40:255–260, 1985.

[81] P. S. Burge and M. N. Richardson. Occupational asthma due to indirect exposure to lauryl dimethyl benzyl ammonium chloride used in a floor cleaner. *Thorax*, 49:842–843, 1994.

[82] A. Innocenti. Asma professionale da benzalconio cloruro. *Med. Lavoro*, 69(6):713–715, 1978.

[83] T. Charles. Asthma due to industrial use of chloramine [letter]. *British Medical Journal*, 2:334, 1979.

[84] M. Bourne, M. Flindt, and J. Walker. Asthma due to industrial use of chloramine. *British Medical Journal*, 2:10–12, 1979.

[85] M Chan-Yeung and J. L. Malo. *Asthma in the Workplace*, chapter Compendium II. Marcel Dekker, New York, 1993.

[86] E. R. Waclawski, L. G. McAlpine, and N. C. Thompson. Occupational asthma in nurses caused by chlorhexidine and alcohol aerosols. *British Medical Journal*, 298:929–930, 1989.

[87] L. Nagy and M. Orosz. Occupational asthma due to hexchlorophene. *Thorax*, 39:630–631, 1984.

[88] D. J. Hendrick, M. J. Connolly, S. C. Stenton, A. G. Bird, I. S. Winterton, and E. H. Walters. Occupational asthma due to sodium iso-nonanoyl oxybenzene sulphonate, a newly developed detergent ingredient. *thorax*, 43:501–502, 1988.

[89] B. Savonius, H. Keskinen, M. Tuppurainen, and L. Kanerva. Occupational asthma caused by ethanolamines. *Allergy*, 49:877–881, 1994.

[90] G. M. Sterling. Asthma due to aluminium soldering flux. *Thorax*, 22:533–537, 1967.

[91] H. Gelfand. Respiratory allergy due to chemical compounds encountered in the rubber, lacquer, shellac, and beauty culture industries. *J Allergy*, 34:374–381, 1963.

[92] J. Pepys and C. Pickering. Asthma due to inhaled chemical fumes - amino-ethyl ethanolamine in aluminium soldering flux. *Clinical Allergy*, 2:197–204, 1972.

[93] M. E. Gadon, J. M. Melius, G. J. McDonald, and D. Orgel. New-onset asthma after exposure to the steam system additive 2-diethylaminoethanol. *Journal of Occupational Medicine*, 36(6):623–626, 1994.

[94] R. E. Brubaker, D. B. Muranko, M.P.H. Smith, Beck G. J., and G. Scovel. Evaluation and control of a respiratory exposure to 3-(dimethylamino)propylamine. *Journal of Occupational Medicine*, 21(10):688–690, 1979.

[95] F. D. Aldrich, A. W. Stange, and R. E. Geesaman. Smoking and ethylene diamine sensitization in an industrial population. *J Occup Med*, 29:311–314, 1987.

[96] S. Lam and M. Chan-Yeung. Ethylenediamine-induced asthma. *Am. Rev. Resp. Dis.*, 121:151–155, 1980.

[97] L. Hagmar, T. Bellander, B. Bergoo, and B. G. Simonsson. Piperazine-induced occupational asthma. *Journal of Occupational Medicine*, 24:193–197, 1982.

[98] J. Pepys, C. A. C. Pickering, and H. W. G. Loudon. Asthma due to inhaled chemical agents – piperazine dihydrochloride. *Clinical Allergy*, 2:189–196, 1972.

[99] L. Belin, U. Wass, G. Audunsson, and L. Mathiasson. Amines: possible causative agents in the development of bronchial hyperreactivity in workers manufacturing polyurethanes from isocyanates. *British Journal of Industrial Medicine*, 40:251–257, 1983.

[100] D. E. Silberman and H. A. Sorrel. Allergy in fur workers with special reference to paraphenylenediamine. *J. Allergy*, 30:11–18, 1959.

[101] E. M. Lambourn, J. P. Hayes, W. A. McAllister, and A. J. New man Taylor. Occupational asthma due to EPO 60. *Br. J. Ind. Med.*, 49:294–295, 1992.

[102] W. Jedrychowski. Styrene and methyl methacrylate in the industrial environment as a risk factor of chronic obstructive lung disease. *Int. Arch. Occup. Environ. Health*, 51:151–157, 1982.

[103] S. Lozewicz, A. G. Davison, A. Hopkirk, P. Burge Sherwood, D. A. Boldy, J. F. Riordan, B. W. Platts, D. Davies, and A. J. Newman-Taylor. Occupational asthma due to methyl methacrylate and cyanoacrylates. *Thorax*, 40:836–839, 1985.

[104] R. M. C. L. Niven. Occupational asthma in workers exposed to cyanoacrylates. *Thorax*, 49:400P, 1994.

[105] C. A. C. Pickering, D. Bainbridge, I. H. Birtwistle, and D. L. Griffiths. Occupational asthma due to methyl methacrylate in an orthopaedic theatre nurse. *British Medical Journal*, 292:1362–1363, 1986.

[106] T. Nakazawa. Occupational asthma due to alkyl cyanoacrylate. *Journal of Occupational Medicine*, pages 709–710, 1990.

[107] C. C. Chan, T. H. Cheong, H. S. Lee, Y. T. Wang, and S. C. Poh. Case of occupational asthma due to glue containing cyanoacrylate. *Annals of the Academy Medicine Singapore*, 23:731–733, 1994.

[108] B. Savonius, H. Keskinen, M. Tuppurainen, and L. Kanerva. Occupational respiratory disease caused by acrylates. *Clinical and Experimental Allergy*, 23:416–424, 1993.

[109] I. Honda, H. Kohrogi, M. Ando, S. Araki, T. Ueno, M. Futatsuka, and A. Ueda. Occupational asthma induced by the fungicide tetrachloro-isophthalonitrile. *Thorax*, 47:760–761, 1992.

[110] S. Royce, P. Wald, D. Sheppard, and J. Balmes. Occupational asthma in a pesticides manufacturing worker. *Chest*, 103:295–296, 1993.

[111] A Weiner. Bronchial asthma due to the organic phosphate insecticides. *Annals of Allergy*, 19:397–401, 1961.

[112] D. Bryant. Asthma due to insecticide sensitivity. *Aust. N. Z. J. Med.*, 15:66–68, 1985.

[113] R. J. Davies, D. J. Hendrick, and J. Pepys. Asthma due to inhaled chemical agents: ampicillin, benzyl penicillin, 6-amino penicillamic acid and related substances. *Clinical Allergy*, 4:227–247, 1974.

[114] G. Briatico-Vangosa, F. Beretta, S. Bianchi, A. Cardani, M. Marchisio, C. Nava, and F. Talamo. Bronchial asthma due to 7-aminocephalosporanic acid (7-ACA) in workers employed in cephalosporine production. [italian]. *Med. Lavoro*, 72:488–493, 1981.

[115] G. Carlesi, E. Ferrea, C. Melino, A. Messineo, and E. Pacelli. Aspects of environmental health and of pathology caused by pollution with amoxicillin in a pharmaceutical industry. *Nuovi Annali D Igiene E Microbiologia*, 30:185–196, 1979.

[116] R. J. Davies and J Pepys. Asthma due to inhaled chemical agents – the macrolide antibiotic spiramycin. *Clinical Allergy*, 1:99–107, 1975.

[117] P. L. Paggiaro, A. M. Loi, and G. Toma. Bronchial asthma and dermatitis due to spiramycin in a chick breeder. *Clinical Allergy*, 9:571–574, 1979.

[118] J. L. Malo and A. Cartier. Occupational asthma in workers of a pharmaceutical company processing spiramycin. *Thorax*, 43:371–377, 1988.

[119] M. Menon and A. Das. Tetracycline asthma - a case report. *Clinical Allergy*, 7:285–290, 1977.

[120] M. G. Harries, A. Newman Taylor, J. Wooden, and A. MacAuslan. Bronchial asthma due to alpha methyldopa. *British Medical Journal*, 1:1461, 1979.

[121] I. Coutts, S. Lozewicz, M. Dally, A. J. Newman-Taylor, P. S. Burge, A. C. Flind, and D. J. H. Rogers. Respiratory symptoms related to work in a factory manufacturing cimetidine tablets. *British Medical Journal*, 288:1418, 1984.

[122] R. M. Agius. Opiate inhalation and occupational asthma. *British Medical Journal*, 298:323, 1989.

[123] R. G. Edwards, J. M. Dewdney, R. J. Dobrzanski, and D. Lee. Immunogenicity and allergenicity studies on two beta-lactam structures, a clavam, clavulanic acid, and a carbapenem: Structure activity relationships. *Int.Archs.Allergy Appl.Immun*, 85:184–189, 1988.

[124] G. Brunetti and G. Moscato. Bronchial asthma due to occupational exposure to dioctyl phthalate [original italian title: Asma bronchiale da esposizione professionale a dioctiftalato]. *La Medicina del Lavoro*, 75(2):120–124, 1984.

[125] J. Kammermeyer and K. Mathews. Hypersensitivity to phenylglycine acid chloride. *J Allergy Clin Immunol*, 52:73–84, 1973.

[126] B. Perrin, J. L. Malo, A. Cartier, S Evans, and J. Dolovich. Occupational asthma in a pharmaceutical work exposed to hydralazine. *Thorax*, 45:980–981, 1990.

[127] M. Rosberg. Asthma bronchiale caused by sulfathiazole. *Acta Medica Scandinavica*, CXXVI:185–190, 1946.

[128] S. Asai, T. Shimoda, K. Hara, and K. Fujiwara. Occupational asthma caused by isonicotinic acid hydrazide (INH) inhalation. *Journal of Allergy and Clinical Immunology*, 80:578–582, 1987.

[129] F. Lagier, A. Cartier, J. Dolovich, and J. L. Malo. occupational asthma in apharmaceutical worker exposed to penicilamine. *Thorax*, 44:157–158, 1989.

[130] I. Fawcett, J. Pepys, and M Erooga. Asthma due to glycyl compound powder - an intermediate in production of salbutamol. *Clinical Allergy*, 6:405–409, 1976.

[131] R. S. Schwettman and C. L. Casterline. Delayed asthmatic response following exposure to enflurane. *Anesthesiology*, 44:166–169, 1976.

[132] R. M. Agius, A. G. Davison, E. R. Hawkins, and A. J. Newman-Taylor. Occupational asthma in salbutamol process workers. *Occupational and Environmental Medicine*, 51:397–399, 1994.

[133] M. Rosenberg, D. Aaronson, and C. Evans. Asthmatic responses to inhaled aminophylline: A report of two cases. *Annals of Allergy*, 52:97–98, 1984.

[134] G. Moscato, P. Marracini, A. Dellabianca, G Vinci, and S. M. Candura. Asma e rinite professionali da stirene. *G. Ital. Med. Lav.*, 10:253–259, 1988.

[135] J. P. Hayes, L. Lambourn, J. A. Hopkirk, S. R. Durham, and A. J. Newman Taylor. Occupational asthma due to styrene. *Thorax*, 46:397–397, 1991.

[136] S. J. Bourke, P. Sandhu, S. C. Stenton, M. Mohan, and D. J. Hendrick. Asthma in roof bolting coal miners. *Thorax*, 49:400P, 1994.

[137] A. Seaton, B. Cherri, and J. Turnbull. Rubber glove asthma. *British Medical Journal*, 296:531–532, 1988.

[138] J. L. Malo, L. Pineau, and A. Cartier. Occupational asthma due to azobisformamide. *Clinical Allergy*, 15:261–264, 1985.

[139] J. C. Normand, F. Grange, C. Hernandez, A. Ganay, P. Davezies, and A. Bergeret. Occupational asthma after exposure to azodicarbonamide: report of four cases. *British Journal of Industrial Medicine*, 46:60–62, 1989.

[140] W. V. Evans and A. Seaton. Hypersensitivity pneumonitis in a technician using pauli's reagent. *Thorax*, 34:767–770, 1979.

[141] P. S. Burge, M. Hendy, and E. S. Hodgson. Occupational asthma, rhinitis and dermatitis due to tetrazene in a detonator manufacturer. *Thorax*, 39:470–471, 1984.

[142] D. W. Cockcroft, A. Cartier, G. Jones, S. M. Tarlo, J. Dolovich, and F. E. Hargreave. Asthma cused by occupational exposure to a furan-based binder system. *Journal of Allergy and Clinical Immunology*, 66(6):458–463, 1980.

[143] S. A. Greene and S. Freedman. Asthma due to inhaled chemical agents – amprolium hydrochloride. *Clinical Allergy*, 6:105–108, 1976.

[144] D. Deschamps, N. Rosenberg, P. Soler, G. Maillard, E. Fournier, D. Salson, and P. Gervais. Persistent asthma after accidental exposure to ethylene oxide. *British Journal of Industrial Medicine*, 49:523–525, 1992.

[145] P. Dugue, C. Faraut, M. Figueredo, A Bettendorf, and J. M. Salvadori. Asthme professionel *à* l'oxyde d'*é*thylène chez une *infirmière*. *La Presse Médicale*, 28:1455, 1991.

[146] D. Choudat, F. Neukirch, P. Brochard, G. Barrat, J. Marsac, F. Conso, and M. Philbert. Allergy and occupational exposure to hydroquinone and methionine. *British Journal of Industrial Medicine*, 45:376–380, 1988.

[147] I. R. White. Occupation dermatitis. *British Medical Journal*, 313:487–489, 1996.

[148] I. Kimber. Contact and respiratory sensitization by chemical allergens: Uneasy relationships. *American Journal of Contact Dermatitis*, 6:34–39, 1995.

[149] I. Kimber, I. L. Berstein, M. H. Karol, M. K. Robinson, K. Sarlo, and M. K. Selgrade. Workshop Overview identification of respiratory allergens. *Fundamental and Applied Toxicology*, 33:1–10, 1996.

[150] Cusano F. and Luciano S. Contact allergy to benzalkonium chloride and glutaraldehyde in a dental nurse. *Contact Dermatitis*, 28(2):127, 1993.

[151] Nethercott J. R., Holness D. L., and Page E. Occupational contact dermatitis due to glutaraldehyde in health care workers. *Contact Dermatitis*, 18(4):193–196, 1988.

[152] Hansen K. S. Occupational dermatoses in hospital cleaning women. *Contact Dermatitis*, 9(5):343–351, 1983.

[153] Estlander T., Keskinen H., Jolanki R., and Kanerva L. Occupational dermatitis from exposure to polyurethane chemicals. *Contact Dermatitis*, 27(3):161–165, 1992.

[154] Wilkinson S. M., Cartwright P. H., Armitage J., and English J. S. Allergic contact dermatitis from 1,6-diisocyanatohexane in an anti-pill finish. *Contact Dermatitis*, 25(2):94–96, 1991.

[155] Bruynzeel D. P. and van der Wegen-Keijser M. H. Contact dermatitis in a cast technician. *Contact Dermatitis*, 28(3):193–194, 1993.

[156] Liden C. Allergic contact dermatitis from 4,4'-diisocyanatoldiphenyl methane (MDI) in a molder. *Contact Dermatitis*, 6(4):301–302, 1980.

[157] English J. S., Foulds I., White I. R., and Rycroft R. J. Allergic contact sensitization to the glycidyl ester of hexahydrophthalic acid in a cutting oil. *Contact Dermatitis*, 15(2):66–68, 1986.

[158] D. Bawden. Computerized chemical structure-handling techniques in structure-activity studies and molecular property prediction. *Journal of Chemical Information and Computer Science*, 23:14–22, 1983.

[159] A. Dalby, J. G. Nourse, W. D. Hounshell, A. K. I. Gushurst, D. L. Grier, B. A. Leland, and J. Lauffer. Description of several chemical structure file formats used by computer programs developed at molecular design limited. *Journal of Chemical Information And Computer Science*, 32:244–255, 1992.

[160] A. R. Katritzky, V. S. Lobanov, and M. Karelson. Short course in heterocyclic chemistry. Privately, Centre for Heterocyclic Compounds, Dept. of Chemistry, University of Florida, 1995.

[161] Health and Safety Executive. *EH40/94 Occupational exposure limits 1994*. HMSO, UK, 1994.

[162] A. C. De Groot. *Patch Testing*. Elsevier, Amsterdam, first edition edition, 1986.

[163] H. L. Keynes, D. J. Ross, and J. C McDonald. SWORD'95: Surveillance of work-related and occupational respiratory disease in the UK. *Occupational Medicine*, 46:379–381, 1996.

[164] R. M. Agius, R. A. Elton, L. Sawyer, and P. Taylor. Occupational asthma and the chemical properties of low molecular mass organic substances. *Occupational Medicine*, 44:34–36, 1994.

[165] G. W. Adamson and D. Bawden. Comparison of hierarchical cluster analysis techniques for automatic classification of chemical structures. *Journal of Chemical Information and Computer Science*, 21:204–209, 1981.

[166] G. M Downs, P. Willett, and W Fisanick. Similarity searching and clustering of chemical-structure databases using molecular property data. *Journal of Chemical Information and Computer Science*, 34:1094–1102, 1994.

[167] R. A. Jarvis and E. A. Patrick. Clustering using a similarity measure based on shared nearest neighbors. *IEEE Transactions on Computers*, C-22(11):1025–1034, 1973.

[168] EPIDECON - course on veterinary epidemiology. Zaragoza, 1995. V1, 16.16, Test Agreement.

[169] G. Eason, C. W. Coles, and G. Gettinby. *Mathematics and Statistics for the Bio-Sciences*. Ellis Horwood, Chichester, 1989.

[170] K Venables and M. Chan-Yeung. Occupational asthma. *The Lancet*, 349:1465–1469, 1997.

[171] I. Kimber and R.J. Dearman. Validation of predictive tests for chemical respiratory allergens. Technical report, Zeneca Central Toxicology laboratory for the Health and Safety Executive, 1998.

[172] P. S. Burge. *Clinics in Immunology and Allergy*, volume 4, chapter Occupational Asthma, rhinitis and alveolitis due to colophony, pages 55–81. WB Saunders and Co., London, U.K., 1984.

[173] M. Bernstein, I.L. Chan-Yeung, J. Malo, and D. I. Bernstein, editors. *Asthma in the Workplace*. New York, Marcel Dekker, New York, 1993.

[174] M. Ben-Ari. *Mathematical Logic for Computer Science*. Prentice Hall, 1993.

[175] J. P. Hayes and A. J. Newman-Taylor. In vivo models of occupational asthma due to low molecular weight chemicals. *Occupational & Environmental Medicine*, 52:539–543, 1995.

[176] R. J. Dearman, D. A. Basketter, and I. Kimber. Characterization of chemical allergens as a function of divergent cytokine secretion profiles induced in mice. *Toxicology and Applied Pharmacology*, 138:308–316, 1996.

[177] M. H. Karol, C. Graham, R. Grealy, O. T. Macina, N. Sussman, and H. S. Rosenkranz. Structure-activity relationships and computer-assisted analysis of respiratory sensitisation potential. *Toxicology Letters*, 86:187–191, 1996.

[178] M. P. Payne, P. T. Walsh, and J. J. Langowski. Skin sensitisation structure-activity relationships for phenols and anilins and application of a qualitative rule-based system:DEREK. In *Proc 9th Europ Symp*, pages 504–506, 1992.

[179] C. Graham, R. Grealy, O. T. Macina, M. H. Karol, and H. S. Rosenkranz. QSAR for allergic contact dermatitis. *Quantitative Structure Activity Relationships*, 15:224–229, 1996.

[180] B.T. Butcher, C.E. O'Neil, M.A. Reed, and J.E. Salvaggio. *The Journal of Allergy and Clinical Immunology*, 66:213–216, 1980.

[181] T.P. Ng, H.S. Lee, Y.T. Wang, V.L. Tay, and K.T. Tan. Occupational asthma due to ethylene diamine. *Annals of the Academy of Medicine*, 20(3):399–402, 1991.

[182] S. Wright and Harman. Ethylenediamine and piperazine sensitivity. *B. M. J.*, 287:463–464, 1983.

[183] H. Welinder and J. Nielsen. Immunologic tests of specific antibodies to organic acid anhydrides. *Allergy*, 4:55–81, 1991.

[184] S. Kivity, E. Fireman, and Y. Lerman. Late asthmatic response to inhaled glacial acetic acid. *Thorax*, 49:727–728, 1994.

[185] R. O. Potts and R. H. Guy. Predicting skin permeability. *Pharmaceutical Research*, 9:663–669, 1992.

[186] S. Menard. *Applied Logistic Regression*, chapter Stepwise Logistic Regression, pages 54–57. Sage University Press, Thousand Oaks, California, 1995.

[187] H. S. Park, M. K. Lee, and C. S. Hong. Reactive dye induced occupational asthma without nonspecific bronchial hyperreactivity. *Yonsei Medical Journal*, 31:98–102, 1990.

[188] E. Epailly, M. Blaumeiser, M. Gonzales, and A. Cantineau. Suspicion d'asthme au chlohydrate de dobutamine. *Therapie*, 49:49–56, 1994.

[189] M. Chan-Yeung, L. MacLean, and P.L. Paggiaro. Follow-up study of 232 patients with occupational asthma caused by western red cedar (thuja plicata). *Journal of Allergy and Clinical Immunology*, 79:792–796, 1987.

[190] I. Fawcett and J. Pepys. Allergy to a tetracycline preparation. *Clinical Allergy*, 6:301–303, 1976.

[191] I.M. O'Brien, A.J. Newman-Taylor, P.S. Burge, M.G. Harries, I.W. Fawcett, and J. Pepys. Toluene di-isocyanate-induced asthma. *Clin. Allergy*, 9:7–15, 1979.

[192] O. Vandenplas, J. L. Malo, M. Saetta, C. E. Mapp, and L. M. Fabbri. Occupational asthma and extrinsic alveolitis due to isocyanates: current status and perspectives. *British Journal of Industrial Medicine*, 50:213–228, 1993.

[193] O. Vandenplas, A. Cartier, J. Lesage, Y. Cloutier, Eng, G. Perreault, L.C. Grammer, M.A. Shaughnessy, and J.L. Malo. Prepolymers of hexamethylene diisocyanate as a cause of occupational asthma. *Journal of Allergy and Clinical Immunology*, 91:850–861, 1993.

[194]

[195] S. Budavari, editor. *The Merck Index*. Merck & Co., Inc., Rahway, N.J., U.S.A., eleventh edition edition, 1989.

[196] Aldrich Chemical Company, The Old Brickyard, New Road, Gillingham, Dorset SP8 4XT. *Catalogue Handbook of Fine Chemicals 1996-1997*.

[197] Bruze M. Kestrup L. Occupational allergic contact dermatitis from diphenylguanidine in a gas mask. *Contact Dermatitis*, 31(2):125–126, 1994.

[198] Patruno C., Suppa F., Sarracco G., and Balato N. Allergic contact dermatitis due to ethyl alcohol. *Contact Dermatitis*, 31(2):124, 1994.

[199] Mathelier-Fusade P. and Leynadier F. Occupational allergic contact reaction to disulfiram. contact dermatitis. *Contact Dermatitis*, 31(2):121–122, 1994.

[200] Steinkraus V. and Hausen B. M. Contact allergy to propylene oxide. *Contact Dermatitis*, 31(2):120, 1994.

[201] Hausen B. M. A case of allergic contact dermatitis due to metanil yellow. *Contact Dermatitis*, 31(2):117–118, 1994.

[202] Goday Bujan JJ., Yanguas Bayona I., and Soloeta Arechavala R. Allergic contact dermatitis from cyanamide: report of 3 cases. *Contact Dermatitis*, 31(2):331–332, 1994.

[203] Conde-Salazar L., Guimaraens D., Romero L., and Harto A. Allergic contact dermatitis to cyanamide (carbodiimide). *Contact Dermatitis*, 7(6):329–330, 1981.

[204] Matthieu L., Weyler J., Deckers I., van Sprundel M., van Andel A., and Dockx P. Occupational contact sensitization to ethylenediamine in a wire-drawing factory. *Contact Dermatitis*, 29(1):39, 1993.

[205] Corazza M., Mantovani L., Trimurti S., and Virgili A. Occupational contact sensitization to ethylenediamine in a nurse. *Contact Dermatitis*, 31(5):328–329, 1994.

[206] Chieregato C., Vincenzi C., Guerra L., and Farina P. Occupational allergic contact dermatitis due to ethylenediamine dihydrochloride and cresyl glycidyl ether in epoxy resin systems. *Contact Dermatitis.*, 30(2):120, 1994.

[207] Dias M., Fernandes C., Pereira F., and Pacheco A. Occupational dermatitis from ethylenediamine. *Contact Dermatitis*, 33(2):129–130, 1995.

[208] Hogan D. J. Allergic contact dermatitis to ethylenediamine. A continuing problem. *Dermatologic Clinics*, 8(1):133–136, 1990.

[209] Ormerod A. D., Wakeel R. A., Mann T. A., Main R. A., and Aldridge R. D. Polyamine sensitization in offshore workers handling drilling muds. *Contact Dermatitis*, 21(5):326–329, 1989.

[210] Bainova A., Khristeva V., Madzhunov I., and Daneva Zh. Dermal exposure to ethylenediamine in the petrochemical industry. [bulgarian title : Dermalno vuzdeistvie na etilendiamina v neftokhimichno proizvodstvo.]. *Problemi na Khigienata*, 12:109–14, 1987.

[211] Crow K. D., Peachey R. D., and Adams J. E. Coolant oil dermatitis due to ethylenediamine. *Contact Dermatitis*, 4(6):359–361, 1978.

[212] Kanerva L., Estlander T., and Jolanki R. Occupational allergic contact dermatitis caused by thiourea compounds. *Contact Dermatitis*, 31(4):242–248, 1994.

[213] Meding B., Baum H., Bruze M., Roupe G., and Trulsson L. Allergic contact dermatitis from diphenylthiourea in vulkan heat retainers. *Contact Dermatitis*, 22(1):8–12, 1990.

[214] Dal Monte A., Laffi G., and Mancini G. Occupational contact dermatitis due to spectinomycin. *Contact Dermatitis*, 31(3):204–205, 1994.

[215] Caraffini S., Assalve D., Stingeni L., and Lisi P. Tylosin, an airborne contact allergen in veterinarians. *Contact Dermatitis*, 31(5):327–328, 1994.

[216] Danese P., Zanca A., and Bertazzoni M. G. Occupational contact dermatitis from tylosin. *Contact Dermatitis*, 30(2):122–123, 1994.

[217] Barbera E. and de la Cuadra J. Occupational airborne allergic contact dermatitis from tylosin. *Contact Dermatitis*, 20(4):308–309, 1989.

[218] Hjorth N. and Roed-Petersen J. Allergic contact dermatitis in veterinary surgeons. *Contact Dermatitis*, 6(1):27–29, 1980.

[219] Ueda A., Aoyama K., Manda F., Ueda T., and Kawahara Y. Delayed-type allergenicity of triforine (saprol). *Contact Dermatitis*, 31(3):140–145, 1994.

[220] Pentel M. T., Andreozzi R. J., and Marks J. G. Jr. Allergic contact dermatitis from the herbicides trifluralin and benefin. *Journal of the American Academy of Dermatology*, 31(6):1057–1058, 1994.

[221] Aguirre A., Manzano D., Zabala R., Raton J. A., and Diaz Perez J. L. Contact allergy to captan in a hairdresser. *Contact Dermatitis*, 31(1):46, 1994.

[222] Vilaplana J. and Romaguera C. Captan, a rare contact sensitizer in hairdressing. *Contact Dermatitis*, 29(2):107, 1993.

[223] Jung H. D., Honemann W., Kloth C., Lubbe D., Pambor M., QuednowC., Ratz K. H., Rothe A., and Tarnick M. Contact eczema

caused by pesticides in east germany. [german title: Kontaktekzem durch pestizide in der deutschen demokratischen republik.]. *Dermatologische Monatsschrift*, 175(4):203–214, 1989.

[224] Lisi P., Caraffini S., and Assalve D. Irritation and sensitization potential of pesticides. *Contact Dermatitis.*, 17(4):212–218, 1987.

[225] Guo Y. L., Wang B. J., Lee J. Y., and Chou S. Y. Occupational hand dermatoses of hairdressers in tainan city. *Occupational & Environmental Medicine*, 51(10):689–692, 1994.

[226] Foti C., Vena GA., and Angelini G. Occupational contact allergy to benzydamine hydrochloride. *Contact Dermatitis*, 27(5):328–329, 1992.

[227] Rademaker M. Allergic contact dermatitis from lavender fragrance in difflam gel. *Contact Dermatitis*, 31(1):58–59, 1994.

[228] Galindo P. A., Garcia R., Garrido J. A., Feo F., and Fernandez F. Allergic contact dermatitis from colour developers: absence of cross-sensitivity to para-amino compounds. *Contact Dermatitis*, 30(5), 1994.

[229] Ibbotson S. H. and Lawrence C. M. Allergic contact dermatitis from aziridine crosslinker cx100. *Contact Dermatitis*, 30:301, 1994.

[230] Garabrant D. H. Dermatitis from aziridine hardener in printing ink. *Contact Dermatitis*, 12:209–212, 1985.

[231] Valsecchi R., Leighissa P., Piazzolla S., Naldi L., and Cainelli T. Occupational contact dermatitis from propranolol. *Contact Dermatitis*, 30(3):177, 1994.

[232] Rudzki E. Occupational dermatitis among health service workers. *Dermatosen in Beruf und Umwelt*, 27(4):112–115, 1979.

[233] Meding B., Barregard L., and Marcus K. Hand eczema in car mechanics. *Contact Dermatitis*, 30(3):129–134, 1994.

[234] Matsunaga K., Hosokawa K., Suzuki M., Arima Y., and Hayakawa R. Occupational allergic contact dermatitis in beauticians. *Contact Dermatitis.*, 18(2):94–96, 1988.

[235] Rycroft R. J. Allergic contact dermatitis from dipentene in honing oil. *Contact Dermatitis*, 6(5):325–329, 1980.

[236] Hansson C. Allergic contact dermatitis from N-(1,3-dimethylbutyl)-N'-phenyl-p-phenylenediamine and from compounds in polymerized 2,2,4-trimethyl-1,2-dihydroquinoline. *Contact Dermatitis*, 30(2):114–115, 1994.

[237] Herve-Bazin B., Gradiski D., Duprat P., Marignac B., Foussereau J., Cavelier C., and Bieber P. Occupational eczema from N-isopropyl -N'-phenylparaphenylenediamine (IPPD) and N-dimethy-1,3 butyl-N'-phenylparaphenylenediamine (DMPPD) in tyres. *Contact Dermatitis*, 3(1):1–15, 1977.

[238] R. Jolanki, L. Kanerva, T. Estlander, and K. Tarvainen. Concomitant sensitization to triglycidyl isocyanurate, diaminodiphenylmethane and 2-hydroxyethyl methacrylate from silk-screen printing coatings in the manufacture of circuit boards. *Contact Dermatitis*, 30:12–15, 1994.

[239] McFadden J. P. and Rycroft R. J. Occupational contact dermatitis from triglycidyl isocyanurate in a powder paint sprayer. *Contact Dermatitis*, 28(4):251, 1993.

[240] Foulds I. S. and Koh D. Allergic contact dermatitis from resin hardeners during the manufacture of thermosetting coating paints. *Contact Dermatitis*, 26(2):87–90, 1992.

[241] Mathias C. G. Allergic contact dermatitis from triglycidyl isocyanurate in polyester paint pigments. *Contact Dermatitis*, 19(1):67–68, 1988.

[242] LeVine M. J. Occupational photosensitivity to diaminodiphenylmethane. *Contact Dermatitis*, 9(6):488–490, 1983.

[243] Van Joost T., Heule F., and de Boer J. Sensitization to methylenedianiline and para-structures. *Contact Dermatitis*, 16(5):246–248, 1987.

[244] Tobler M., Wuthrich B., and Freiburghaus A. U. Contact dermatitis from acrylate and methacrylate compounds in lowicryl embedding media for electron microscopy. *Contact Dermatitis*, 23(2):96–102, 1990.

[245] Pedersen NB., Senning A., and Nielsen A. O. Different sensitising acrylic monomers in napp printing plate. *Contact Dermatitis*, 9(6):459–464, 1983.

[246] Mathias C. G., Caldwell T. M., and Maibach H. I. Contact dermatitis and gastrointestinal symptoms from hydroxyethylmethacrylate. *British Journal of Dermatology*, 100(4):447–449, 1979.

[247] Bork K. Allergic contact dermatitis on a violinist's neck from para-phenylenediamine in a chin rest stain. *Contact Dermatitis*, 28(4):250–251, 1993.

[248] Guerra L., Tosti A., Bardazzi F., Pigatto P., Lisi P., Santucci B., Valsecchi R., Schena D., Angelini G., and Sertoli A. et al. Contact dermatitis in hairdressers: the italian experience. *Contact Dermatitis*, 26(2):101–107, 1992.

[249] Fowler J. F. Jr. Occupational dermatitis from stamp pad ink. *Contact Dermatitis*, 16(1):38, 1987.

[250] Nethercott J. R., MacPherson M., Choi B. C., and Nixon P. Contact dermatitis in hairdressers. *Contact Dermatitis*, 14(2):73–79, 1986.

[251] Liden C. and Brehmer-Andersson E. Occupational dermatoses from colour developing agents. clinical and histopathological observations. *Acta Dermato-Venereologica*, 68(6):514–522, 1988.

[252] Edwards E. K Jr. and Edwards E. K. Contact urticaria and allergic contact dermatitis caused by paraphenylenediamine. *Cutis*, 34(1):87–88, 1984.

[253] Kanerva L., Jolanki R., and Estlander T. Dentist's occupational allergic contact dermatitis caused by coconut diethanolamide, N-ethyl-4-toluene sulfonamide and 4-tolyldiethanolamine. *Acta Dermato-Venereologica*, 73(2):126–129, 1993.

[254] Guerra L., Vincenzi C., Bardazzi F., and Tosti A. Contact sensitization due to isophoronediamine. *Contact Dermatitis*, 27(1):52–53, 1992.

[255] Serra-Baldrich E. and Camarasa J. G. Contact dermatitis caused by isophorone diamine. [spanish title: Dermatitis de contacto por isoforendiamina (IPD).]. *Medicina Cutanea Ibero-Latino-Americana*, 17(3):175–177, 1989.

[256] Lachapelle J. M., Tennstedt D., and Dumont-Fruytier M. Occupational allergic contact dermatitis to isophorone diamine (IPD) used as an epoxy resin hardener. *Contact Dermatitis*, 4(2):109–112, 1978.

[257] Watsky K. L., Reynolds K., Berube D., and Bayer F. J. Occupational contact dermatitis from tosyl chloride in a chemist. *Contact Dermatitis*, 29(4):211–212, 1993.

[258] Schmidt R. J. Tosyl chloride [letter; comment]. *Contact Dermatitis*, 32(2):124–125, 1995.

[259] Quirce S., Olaguibel J. M., Garcia B. E., and Tabar A. I. Occupational airborne contact dermatitis due to benzoyl peroxide. *Contact Dermatitis*, 29(3):165–166, 1993.

[260] Bonnekoh B. and Merk H. F. Airborne allergic contact dermatitis from benzoyl peroxide as a bleaching agent of candle wax. *Contact Dermatitis*, 24(5):367–368, 1991.

[261] Fisher A. A. Allergic bakers' dermatitis due to benzoyl peroxide [news]. *Cutis*, 43(2):128–129, 1989.

[262] Kanerva L., Estlander T., Jolanki R., and Henriks-Eckerman M. L. Occupational allergic contact dermatitis caused by diethylenetriamine in carbonless copy paper. *Contact Dermatitis*, 29(3):147–151, 1993.

[263] Meding B. Allergic contact dermatitis from diethylenetriamine in a goldsmith workshop. *Contact Dermatitis*, 8(2):142, 1982.

[264] McFadden J. P., Kinoulty M., and Rycroft R. J. Allergic contact dermatitis from the fungicide bupirimate. *Contact Dermatitis*, 28(1):47, 1993.

[265] Geier J. and Fuchs T. Contact allergy due to 4-N,N-dimethylaminobenzene diazonium chloride and thiourea in diazo copy paper. *Contact Dermatitis*, 28(5):304–305, 1993.

[266] Dooms-Goossens A., Chrispeels M. T., De Veylder H., Roelandts R., and Willems L.and Degreef H. Contact and photocontact sensitivity problems associated with thiourea and its derivatives: a review of the literature and case reports. *British Journal of Dermatology*, 116(4):573–579, 1987.

[267] Kanerva L., Estlander T., Jolanki R., and Tarvainen K. Occupational allergic contact dermatitis caused by exposure to acrylates during work with dental prostheses. *Contact Dermatitis*, 28(5):268–275, 1993.

[268] Kanerva L. and Verkkala E. Electron microscopy and immunohistochemistry of toxic and allergic effects of methylmethacrylate on the skin. *Archives of Toxicology. Supplement.*, 9:456–459, 1986.

[269] Hansen L., Hammershoy O., and Boll P. M. Allergic contact dermatitis from falcarinol isolated from schefflera arboricola. *Contact Dermatitis*, 14(2):91–93, 1986.

[270] Reid C. M. and Rycroft R. J. Allergic contact dermatitis from multiple sources of MCI/MI biocide and formaldehyde in a printer. *Contact Dermatitis*, 28(4):252–253, 1993.

[271] Rycroft R. J. and Neild V. S. Allergic contact dermatitis from MCI/MI biocide in a printer. *Contact Dermatitis*, 26(2):142, 1992.

[272] Pilger C., Nethercott J. R., and Weksberg F. Allergic contact dermatitis due to a biocide containing 5-chloro-2-methyl-4-isothiazolin-3-one. *Contact Dermatitis*, 14(4):201–204, 1986.

[273] Foussereau J., Brandle I., and Boujnah-Khouadja A. Allergic contact eczema caused by isothiazolin-3-one derivatives. [german title : Allergisches kontaktekzem durch isothiazolin-3-onderivate.]. *Dermatosen in Beruf und Umwelt*, 32(6):208–211, 1984.

[274] Knegt-Junk C., Geursen-Reitsma L., and van Joost T. Allergic contact dermatitis from pyridine in karl fischer reagent. *Contact Dermatitis.*, 28(4):252, 1993.

[275] Holness D. L. and Nethercott J. R. The performance of specialized collections of bisphenol A epoxy resin system components in the evaluation of workers in an occupational health clinic population. *Contact Dermatitis*, 28(4):216–219, 1993.

[276] Tomb R. R., Lepoittevin J. P., Durepaire F., and Grosshans E. Ectopic contact dermatitis from ethyl cyanoacrylate instant adhesives. *Contact Dermatitis*, 28(4):206–208, 1993.

[277] Corazza M. and Virgili A. Airborne allergic contact dermatitis from benzalkonium chloride. *Contact Dermatitis*, 28(3):195–196, 1993.

[278] Klein G. F., Sepp N., and Fritsch P. Allergic reactions to benzalkonium chloride? do the use test. *Contact Dermatitis*, 25(4):269–270, 1991.

[279] Kawai K., Nakagawa M., Sasaki Y., and Kawai K. Occupational contact dermatitis from kathon 930. *Contact Dermatitis*, 28(2):117–118, 1993.

[280] Speight E. L., Beck M. H., and Lawrence C. M. Occupational allergic contact dermatitis due to 3-dimethylaminopropylamine. *Contact Dermatitis*, 28(1):49–50, 1993.

[281] Won J. H., Ahn S. K., and Kim S. C. Allergic contact dermatitis from the herbicide alachlor. *Contact Dermatitis*, 28(1):38–39, 1993.

[282] Oleaga J. M., Aguirre A., Landa N., Gonzalez M., and Diaz-Perez J. L. Allergic contact dermatitis from kathon 893. *Contact Dermatitis*, 27(5):345–346, 1992.

[283] Mathias C. G., Andersen K. E., and Hamann K. Allergic contact dermatitis from 2-n-octyl-4-isothiazolin-3-one, a paint mildewcide. *Contact Dermatitis*, 9(6):507–509, 1983.

[284] Aguirre A., Landa N., Gonzalez M., and Diaz-Perez J. L. Allergic contact dermatitis in a photographer. *Contact Dermatitis*, 27(5):340–341, 1992.

[285] Nater J. P. Hypersensitivity to rubber. [original german title : Uberempfindlichkeit gegen gummi]. *Berufs-Dermatosen*, 23(5):161–168, 1975.

[286] Hausen B. M. Occupational allergic contact dermatitis from ethyl chloro oximido acetate. *Contact Dermatitis*, 27(4):277–278, 1992.

[287] Sanz-Gallen P., Planas J., Martinez P., and Gimenez-Arnau J. M. Allergic contact dermatitis due to 1,2-benzisothiazolin-3-one in paint manufacture. *Contact Dermatitis*, 27(4):271–272, 1992.

[288] Dias M., Lamarao P., and Vale T. Occupational contact allergy to 1,2-benzisothiazolin-3-one in the manufacture of air fresheners. *Contact Dermatitis*, 27(3):205–207, 1992.

[289] Damstra R. J., van Vlotten W. A., and van Ginkel C. J. Allergic contact dermatitis from the preservative 1,2-benzisothiazolin-3-one (1,2-BIT; proxel): a case report, its prevalence in those occupationally at risk and in the general dermatological population, and its relationship to allergy to its analogue kathon CG. *Contact Dermatitis*, 27(2):105–109, 1992.

[290] Freeman S. Allergic contact dermatitis due to 1,2-benzisothiazolin-3-one in gum arabic. *Contact Dermatitis*, 11(3):146–149, 1984.

[291] Roberts D. L., Messenger A. G., and Summerly R. Occupational dermatitis due to 1,2-benzisothiazolin-3-one in the pottery industry. *Contact Dermatitis*, 7(3):145–147, 1981.

[292] Fumagalli M., Bigardi A. S., Legori A., and Pigatto P. D. Occupational contact dermatitis from airborne nicergoline. *Contact Dermatitis*, 27(4):256, 1992.

[293] Sanmartin O. and de la Cuadra J. Occupational contact dermatitis from cyclohexanone as a PVC adhesive. *Contact Dermatitis*, 27(3):189–190, 1992.

[294] Hsu C. K., Sun C. C., Su M. S., Kuo E. F., and Wu Y. C. Systemic contact allergy from occupational contact with ethyl ethoxymethylene cyanoacetate. *Contact Dermatitis*, 27(1):58–59, 1992.

[295] O'Driscoll J. B., Marcus R., and Beck M. H. Occupational allergic contact dermatitis from triphenyl phosphite. *Contact Dermatitis*, 20(5):392–393, 1989.

[296] Paulsen E., Andersen K. E., Carlsen L., and Egsgaard H. Carvone: an overlooked contact allergen cross-reacting with sesquiterpene lactones?. *Contact Dermatitis*, 29(3):138–143, 1993.

[297] Meding B., Toren K., Karlberg A. T., Hagberg S., and Wass K. Evaluation of skin symptoms among workers at a swedish paper mill. *American Journal of Industrial Medicine*, 23(5):721–728, 1993.

[298] Wilkinson S. M., Cartwright P. H., and English J. S. Allergic contact dermatitis from mercaptobenzothiazole in a releasing fluid. *Contact Dermatitis*, 23(5):370, 1990.

[299] Hung D. Z., Deng J. F., and Tsai W. J. Dermatitis caused by dimethyl cyanocarbonimidodithioate. *Journal of Toxicology - Clinical Toxicology*, 30(3):351–358, 1992.

[300] Veraldi S., Benelli C., and Pigatto P. D. Occupational allergic contact dermatitis from minoxidil. *Contact Dermatitis*, 26(3):211–212, 1992.

[301] Warin A. P. Allergic contact dermatitis from dazomet. *Contact Dermatitis*, 26(2):135–136, 1992.

[302] Pambor M. and Poweleit H. Allergic contact dermatitis due to diazo copy paper. *Contact Dermatitis*, 26(2):131–132, 1992.

[303] Sengel D., Khelladi A., and Foussereau J. Occupational allergy to diazo-paper used in the textile industry. [original french title : Allergie professionnelle au papier diazo dans l'industrie textile.]. *Dermatosen in Beruf und Umwelt*, 27(6):178–179, 1979.

[304] Laine R., Kanerva L., Tarvainen K., Jolanki R., Estlander T., and Helander I. Nitroglycerin-induced allergic contact dermatitis. [original finnish title: Glyseryylinitraatin aiheuttama allerginen kosketusihottuma. *Duodecim*, 107(1):41–46, 1991.

[305] Kanerva L., Laine R., Jolanki R., Tarvainen K., Estlander T., and Helander I. Occupational allergic contact dermatitis caused by nitroglycerin. *Contact Dermatitis*, 24(5):356–362, 1991.

[306] Brasch J., Hessler H. J., and Christophers E. Occupational (photo)allergic contact dermatitis from azaperone in a piglet dealer. *Contact Dermatitis*, 25(4):258–259, 1991.

[307] A. Ingber. Occupational allergic contact dermatitis from methyl chloroform (1,1,1-trichloroethane). *Contact Dermatitis*, 25(3):193, 1991.

[308] Whitfeld M. and Freeman S. Allergic contact dermatitis to ultra violet cured inks. *Australasian Journal of Dermatology.*, 32(2):65–68, 1991.

[309] Jolanki R. Occupational skin diseases from epoxy compounds. epoxy resin compounds, epoxy acrylates and 2,3-epoxypropyl trimethyl ammonium chloride. *Acta Dermato-Venereologica. Supplementum.*, 159:1–80, 1991.

[310] Kanerva L., Jolanki R., and Estlander T. Hair bulb accumulation of langerhans cells in allergic patch tests. *Acta Dermato-Venereologica. Supplementum.*, 134:64–68, 1987.

[311] Estlander T., Jolanki R., and Kanerva L. Occupational dermatitis to 2,3-epoxypropyl trimethyl ammonium chloride. *Contact Dermatitis*, 14(1):49–52, 1986.

[312] Reygagne A., Garnier R., Efthymiou M. L., and Gervais P. [glycerol monothioglycolate eczema in a hairdresser. persistence of the allergen in the hair several weeks after the application of a permanent]. [french]. *Journal de Toxicologie Clinique et Experimentale*, 11:183–7, 1991.

[313] F. J. Storrs. Permanent wave contact dermatitis: contact allergy to glyceryl monothioglycolate. *Journal of the American Academy of Dermatology*, 11(1):74–85, 1984.

[314] Warshawshki L., Mitchell J. C., and Storrs F. J. Allergic contact dermatitis from glyceryl monothioglycolate in hairdressers. *Contact Dermatitis*, 7(6):351–2, 1981.

[315] Brasch J. Allergic contact dermatitis from 4-chloro-7-nitrobenzofurazan. *Contact Dermatitis*, 25(2):121–124, 1991.

[316] Fisher A. A. Allergic contact dermatitis to mitomycin-C. *Cutis*, 47(4):225–227, 1991.

[317] Kanerva L., Jolanki R., Tupasela O., Halmepuro L., Keskinen H., Estlander T., and Sysilampi ML. Immediate and delayed allergy from epoxy resins based on diglycidyl ether of bisphenol A. *Scandinavian Journal of Work, Environment & Health.*, 17(3):208–215, 1991.

[318] Jolanki R., Kanerva L., Estlander T., Tarvainen K., Keskinen H., and Henriks-Eckerman M. L. Occupational dermatoses from epoxy resin compounds. *Contact Dermatitis*, 23(3):172–183, 1990.

[319] Richter G. and Kadner H. Allergic contact eczema caused by m-xylylene-diamine in the polyurethane silk production. [ger-

man - allergische kontaktekzeme durch m-xylylen-diamin in der polyurethanseidenproduktion.]. *Dermatosen in Beruf und Umwelt*, 38(4):117–120, 1990.

[320] Dooms-Goossens A., Bedert R., Degreef H., and Vandaele M. Airborne allergic contact dermatitis from kitasamycin and midecamycin. *Contact Dermatitis*, 23(2):118–119, 1990.

[321] Mancuso G., Staffa M., Errani A., Berdondini R. M., and Fabbri P. Occupational dermatitis in animal feed mill workers. *Contact Dermatitis*, 22(1):37–41, 1990.

[322] Lo J. S., Taylor J. S., and Oriba H. Occupational allergic contact dermatitis to airborne nitrofurazone. *Dermatologic Clinics*, 8(1):165–168, 1990.

[323] Ancona A. Allergic contact dermatitis to nitrofurazone. 13(1):35, Contact Dermatitis.

[324] Schauder S. The dangers of olaquindox. photoallergy, chronic photosensitive dermatitis and extreme increased photosensitivity in the human, hypoaldosteronism in swine. [german title: Gefahren durch olaquindox. photoallergie, chronisch photosensitive dermatitis und extrem gesteigerte lichtempfindlichkeit beim menschen, hypoaldosteronismus beim schwein.]. *Dermatosen in Beruf und Umwelt*, 37(5):183–185, 1989.

[325] Bedello P. G., Goitre M., Cane D., and Roncarolo G. Allergic contact dermatitis to bayo-N-OX-I. *Contact Dermatitis*, 12(5):284, 1985.

[326] Bruze M. and Almgren G. Occupational dermatoses in workers exposed to epoxy-impregnated fiberglass fabric. *Dermatosen in Beruf und Umwelt*, 37(5):171–176, 1989.

[327] Nethercott J. R. Allergic contact dermatitis due to an epoxy acrylate. *British Journal of Dermatology*, 104(6):697–703, 1981.

[328] Conde-Salazar L., Gonzalez M. A., Guimaraens D., and Romero L. Occupational allergic contact dermatitis from styrene. *Contact Dermatitis*, 21(2):112, 1989.

[329] Gola M., Acciai M. C., Brusi C., Giorgini S., and Sertoli A. Allergic contact dermatitis in a pastry cook. *Contact Dermatitis*, 21(1):57, 1989.

[330] Ingemann Larsen A., Riis Jepsen J., and Thulin H. Allergic contact dermatitis from thiamine. *Contact Dermatitis*, 20(5):387–388, 1989.

[331] Baruffini A., Cirla A. M., Pisati G., Ratti R., and Zedda S. Allergic contact dermatitis from 1,2-dichloropropane. *Contact Dermatitis*, 20(5):379–380, 1989.

[332] Hoffman T. E. and Adams R. M. Contact allergic dermatitis to dicyclohexylcarbodiimide used in protein synthesis. *Journal of the American Academy of Dermatology*, 21(2 Pt 2):436–437, 1989.

[333] Funfstuck V., Knopf B., and Hipler C. Contact allergy to dicyclohexylcarbodiimide. [german title : Kontaktallergie gegenuber dicyclohexylcarbodiimid.]. *Dermatosen in Beruf und Umwelt*, 34(4):110–111, 1986.

[334] Lombardi P., Gola M., Acciai M. C., and Sertoli A. Unusual occupational allergic contact dermatitis in a nurse. *Contact Dermatitis*, 20(4):302–303, 1989.

[335] Dooms-Goossens A., Gevers D., Mertens A., and Vanderheyden D. Allergic contact urticaria due to chloramine. *Contact Dermatitis*, 9(4):319–320, 1983.

[336] van Ketel W. G. and van den Berg W. H. The problem of the sensitization to dithiocarbamates in thiuram-allergic patients. *Dermatologica*, 169(2):70–75, 1984.

[337] Schubert H. Allergic contact dermatitis due to propachlor (author's transl) [german] title : Allergisches kontaktekzem durch propachlor.]. *Dermatologische Monatsschrift.*, 165(7):495–498, 1979.

[338] Carmichael A. J., Foulds I. S., and Sadhra S. Allergic contact dermatitis from tetrafluoroterephthalonitrile and TFX diamine. *Contact Dermatitis*, 20(3):233–234, 1989.

[339] Kanerva L., Estlander T., and Jolanki R. Allergic contact dermatitis from dental composite resins due to aromatic epoxy acrylates and aliphatic acrylates. *Contact Dermatitis*, 20(3):201–211, 1989.

[340] Libow L. F., Ruszkowski A. M., and DeLeo V. A. Allergic contact dermatitis from para-chloro-meta-xylenol in lurosep soap. *Contact Dermatitis*, 20(1):67–68, 1989.

[341] Adams R. M. P-chloro-m-xylenol in cutting fluids: two cases of allergic contact dermatitis in machinists. *Contact Dermatitis*, 7(6):341–343, 1981.

[342] Conde-Salazar L., Llinas M. G., Guimaraens D., and Romero L. Occupational allergic contact dermatitis from amethocaine. *Contact Dermatitis*, 19(1):69–70, 1988.

[343] Senff H., Kuhlwein A., and Hausen B. M. Allergic contact eczema caused by dicyanodiamide]. [german title ; allergisches kontaktekzem auf dicyandiamid.]. *Dermatosen in Beruf und Umwelt*, 36(3):99–101, 1988.

[344] Hayakawa R., Arima Y., Hirose O., and Takeuchi Y. Allergic contact dermatitis due to hexamethylenetetramine in core molding. *Contact Dermatitis*, 18(4):226–228, 1988.

[345] English J. S. and Rycroft R. J. Allergic contact dermatitis from methyl heptine and methyl octine carbonates. *Contact Dermatitis*, 18(3):174–175, 1988.

[346] Noster U. and Hausen B. M. Occupational dermatitis due to a yellow quinophthalone dye (solvent yellow 33: C.I. 47 000)]. [german] [original title : Berufsbedingtes kontaktekzem durch gelben chinophthalonfarbstoff (solvent yellow 33; C.I. 47 000). *Hautarzt*, 29(3):153–157, 1978.

[347] Deschamps D., Garnier R., Savoye J., Chabaux C., Efthymiou M. L., and Fournier. Allergic and irritant contact dermatitis from diethyl-beta-chloroethylamine. *Contact Dermatitis*, 18(2):103–105, 1988.

[348] Shehade S. A., Beck M. H., and Chalmers R. J. Allergic contact dermatitis to crystal violet lactone [corrected] in carbonless copy paper. *Contact Dermatitis*, 17(5):310–311, 1987.

[349] Alomar A., Puig L., and Vilaltella I. Allergic contact dermatitis due to ranitidine. *Contact Dermatitis*, 17(1):54–55, 1987.

[350] Goh C. L. and Ng S. K. Allergic contact dermatitis to ranitidine. *Contact Dermatitis*, 11(4):252, 1984.

[351] Pryce D. W. Allergic contact dermatitis from 2-chlorobenzoyl chloride azine. *Contact Dermatitis*, 16(5):285–6, 1987.

[352] Crijns M. B. and Boom B. W. van der Schroeff J. G. Allergic contact dermatitis to a diazonium compound in copy paper. *Contact Dermatitis*, 16(2):112–113, 1987.

[353] Peters K. and Andersen K. E. Allergic hand dermatitis from 2-hydroxyethyl-acrylate in contact lenses. *Contact Dermatitis*, 15(3):188–189, 1986.

[354] Maurice P. D. and Rycroft R. J. Allergic contact dermatitis from UV-curing acrylate in the manufacture of optical fibres. *Contact Dermatitis*, 15(2):92–93, 1986.

[355] Dahlquist I., Fregert S., and Trulson L. Contact allergy to trimethylolpropane triacrylate (TMPTA) in an aziridine plastic hardener. *Contact Dermatitis*, 9(2):122–124, 1983.

[356] Nethercott J. R., Jakubovic H. R., Pilger C., and Smith J. W. Allergic contact dermatitis due to urethane acrylate in ultraviolet cured inks. *British Journal of Industrial Medicine*, 40(3):241–250, 1983.

[357] Bjorkner B., Dahlquist I., and Fregert S. Allergic contact dermatitis from acrylates in ultraviolet curing inks. *Contact Dermatitis*, 6(6):405–409, 1980.

[358] Emmett E. A. and Kominsky J. R. Allergic contact dermatitis from ultraviolet cured inks. allergic contact sensitization to acrylates. *Journal of Occupational Medicine.*, 19(2):113–115, 1976.

[359] Sonnex T. S. and Rycroft R. J. Dermatitis from phenyl salicylate in safety spectacle frames. *Contact Dermatitis*, 14(5):268–270, 1986.

[360] English J. S., Rycroft R. J., and Calnan C. D. Allergic contact dermatitis from aminotriazole. *Contact Dermatitis*, 14(4):255–256, 1986.

[361] Sonnex T. S. and Rycroft R. J. Allergic contact dermatitis from orthobenzyl parachlorophenol in a drinking glass cleaner. *Contact Dermatitis*, 14(4):247–248, 1986.

[362] Kanerva L., Jolanki R., and Estlander T. Occupational dermatitis due to an epoxy acrylate. *Contact Dermatitis*, 14(2):80–84, 1986.

[363] Valsecchi R., Leghissa P., and Piazzolla S. Tego allergy in the food industry. *Contact Dermatitis*, 23(3):188–189, 1990.

[364] Sinclair S. and Hindson C. Allergic contact dermatitis from dodecyldiaminoethyl glycine. *Contact Dermatitis*, 18(5):320, 1988.

[365] Foussereau J., Samsoen M., and Hecht M. T. *Contact Dermatitis*, 9(3):233–234, 1983.

[366] Suhonen R. Contact allergy to dodecyl-di-(aminoethyl) glycine (desimex i). *Contact Dermatitis*, 6(4):290–291, 1980.

[367] Pfeiff B. Contact dermatitis due to chloroacetophenone (tear gas)]. [german title : Kontaktdermatitis auf chloracetophenon (tranengas).]. *Zeitschrift fur Hautkrankheiten*, 60(1-2):178–80, 183–184, 1985.

[368] Goh C. L. Allergic contact dermatitis to mace tear gas. *Australasian Journal of Dermatology*, 28(3):115–116, 1987.

[369] Fuchs T. and Ippen H. Contact allergy to CN and CS tear gas. [german title : Kontaktallergie auf CN- und CS-tranengas.]. *Dermatosen in Beruf und Umwelt*, 34(1):12–14, 1986.

[370] Goh C. L. Occupational dermatitis from soldering flux among workers in the electronics industry. *Contact Dermatitis*, 13(2):85–90, 1985.

[371] Rebandel P. and Rudzki E. Occupational contact sensitivity in oculists. *Contact Dermatitis*, 15(2):92, 1986.

[372] Falk E. S., Hektoen H., and Thune P. O. Skin and respiratory tract symptoms in veterinary surgeons. *Contact Dermatitis*, 12(5):274–278, 1985.

[373] Fisher A. A. Allergic contact dermatitis to penicillin and streptomycin. *Cutis*, 32(4):314, 318, 324, 1983.

[374] Heine A. and Schulz B. Allergic contact eczema caused by alpha-naphthylisothiocyanate in a chemistry laboratory assistant]. [german title : Allergisches kontaktekzem durch alpha-naphthylsenfol bei einer chemielaborantin.]. *Dermatologische Monatsschrift*, 171(3):201–204, 1985.

[375] Rycroft R. J. Allergic contact dermatitis from a novel diamino intermediate, 5-[(2-aminoethyl)thiomethyl]-N, N-dimethyl-2-furanmethanamine, in laboratory synthesis. *Contact Dermatitis*, 9(6):456–458, 1983.

[376] Yamasaki R., Dekio S., and Jidoi J. Allergic contact dermatitis to ammonium thioglycolate. *Contact Dermatitis*, 11(4):255, 1984.

[377] Nethercott J. R., Gupta S., Rosen C., Enders L. J., and Pilger C. W. Tetraethylene glycol diacrylate. A cause of delayed cutaneous irritant reaction and allergic contact dermatitis. *Journal of Occupational Medicine*, 26(7):513–516, 1984.

[378] Beurey J., Mougeolle J. M., and Weber M. Cutaneous manifestations due to acrylic resins used in printing. [original french title : Accidents cutanes des resines acryliques dans l'imprimerie]. *Annales de Dermatologie et de Syphiligraphie*, 103(4):423–430, 1976.

[379] Kanerva L., Jolanki R., Plosila M., and Estlander T. Contact dermatitis from dibutylthiourea. report of a case with fine structural observations of epicutaneous testing with dibutylthiourea. *Contact Dermatitis*, 10(3):158–162, 1984.

[380] Goh C. L. Allergic contact dermatitis from tetryl and trinitrotoluene. *Contact Dermatitis*, 10(2):108, 1984.

[381] Pedersen N. B., Chevallier M. A., and Senning A. Secondary acrylamides in nyloprint printing plate as a source of contact dermatitis. *Contact Dermatitis*, 8(4):256–262, 1982.

[382] Malten K. E., van der Meer-Roosen C. H., and Seutter E. Nyloprint-sensitive patients react to NN' methylene bis acrylamide. contact dermatitis. 4(4):214–222, 1978.

[383] Beck M. H. and King C. M. Allergic contact dermatitis to epichlorhydrin in a solvent cement. *Contact Dermatitis.*, 9(4):315, 1983.

[384] White I. R., Stewart J. R., and Rycroft R. J. Allergic contact dermatitis from an organic di-isocyanate. *Contact Dermatitis*, 9(4):300–303, 1983.

[385] Hoffman T. E. Allergic contact dermatitis to new plastic resins. *Archives of Dermatology*, 118(12):962, 1982.

[386] Johnsson M., Buhagen M., Leira H. L., and Solvang S. Fungicide-induced contact dermatitis. *Contact Dermatitis*, 9(4):285–288, 1983.

[387] Smith W. D. Allergic dermatitis due to a triacrylate in ultraviolet cured inks. *Contact Dermatitis*, 3(6):312–314, 1977.

[388] Fardal R. W. and Curphey E. R. Phototypesetting paper as a cause of allergic contact dermatitis in newspaper production workers. *Cutis*, 31(5):509–512, 515–517, 1983.

[389] Bruze M. and Fregert S. Allergic contact dermatitis to chloridazon. *Contact Dermatitis*, 8(6):427, 1982.

[390] Kilpikari I. Occupational contact dermatitis among rubber workers. *Contact Dermatitis*, 8(6):359–362, 1982.

[391] Karlsmark T., Weismann K., and Pock-Steen B. Allergic contact eczema following use of sodium methyldithiocarbamate (metan-na) for eradicating roots in drains. [danish title : Allergisk kontakteksem efter anvendelse af natrium-metylditiokarbamat (metam-na) til rodbekaempelse.]. *Ugeskrift for Laeger*, 144(24):1782–1783, 1982.

[392] Adams R. M. Contact allergic dermatitis due to diethylthiourea in a wetsuit. *Contact Dermatitis*, 8(4):277–278, 1982.

[393] Pickering F. C. and Ive F. A. Allergic contact dermatitis from 4,7-dichloroquinoline. *Contact Dermatitis*, 8(4):269–70, 1982.

[394] Carle J. S. and Christophersen C. Dogger bank itch. 4. an eczema-causing sulfoxonium ion from the marine animal, alcyonidium gelatinosum [bryozoa]. *Toxicon.*, 20(1):307–310, 1982.

[395] Grimalt F. and Romaguera C. Cutaneous sensitivity to benzidine. *Dermatosen in Beruf und Umwelt*, 29(4):95–97, 1981.

[396] Rycroft R. J. Allergic contact dermatitis from laboratory synthesis of 4-bromomethyl-6,8-dimethyl-2(1H)-quinolone. *Contact Dermatitis*, 7(1):39–42, 1981.

[397] N Anker and F da Cunha Bang. Long-term intravenous rifampicin treatment. advantages and disadvantages. *European Respiratory Journal*, 62:84–86, 1981.

[398] Stute J., Hausen B. M., and Schulz K. H. Diphenylcyclopropenone - a new strong contact sensitizer (author's transl) [german title : Diphenylcyclopropenon - ein stark wirksames kontaktallergen]. *Dermatosen in Beruf und Umwelt*, 29(1):12–14, 1981.

[399] Shmunes E. and Kempton R. J. Allergic contact dermatitis to dimethoxane in a spin finish. *Contact Dermatitis*, 6(6):421–424, 1980.

[400] Agren S., Dahlquist I., Fregert S., and Persson K. Allergic contact dermatitis from the preservative N-methylol-chloracetamide. contact dermatitis. 6(4):302–303, 1980.

[401] L. Conde-Salazar, A. Garcia Diez, F. Rafeensperger, and Hausen B.M. Contact allergy to the brazilian rosewood substitute machaerium scleroxylon tul. (pao ferro). *Contact Dermatitis*, 6:246–250, 1980.

[402] Richter G. Allergic contact dermatitis from methylisothiocyanate in soil disinfectants. *Contact Dermatitis*, 6(3):183–186, 1980.

[403] Suskind R. R. and Majeti V. A. Occupational and environmental allergic problems of the skin. *Journal of Dermatology*, 3(1):3–12, 1976.

[404] Zschunke E., Rothe A., and Folesky H. Rare allergens - hidden sources of allergens. *Acta Dermato-Venereologica. Supplementum.*, 59(85):207–209, 1979.

[405] Dick D. C. and Adams R. H. Allergic contact dermatitis from monosulfiram (tetmosol) soap. *Contact Dermatitis*, 5(3):199, 1979.

[406] Alfonzo C. Allergic contact dermatitis to isopropylaminodiphenylamine (IPPD). *Contact Dermatitis*, 5(3):145–147, 1979.

[407] Foussereau J. and Cavelier C. Has N-isopropyl-N'-phenylparaphenylenediamine a place among standard allergens? importance of this allergen in rubber intolerance. [original french title : La N-isopropyl'N'-phenylparaphenylenediamine a-t-elle sa place dans la batterie standard d'allergenes? importance de cet allergene dans l'intolerance au caoutchouc.]. *Dermatologica*, 155(3):164–167, 1977.

[408] Fregert S. Allergic contact dermatitis from ethylacrylate in a window sealant. *Contact Dermatitis*, 4(1):56, 1978.

[409] M. Hausen B, H.D. Herrmann, and G. Willuhn. The sensitizing capacity of compositae plants. i. occupational contact dermatitis from arnica longifolia eaton. *Contact Dermatitis*, 4:3–10, 1978.

[410] Bruevich T. S. and Zakharov G. A. Semi-synthetic penicillins, ampicillin and oxacillin and their role in development of occupational allergic dermatoses. [RUSSIAN] [original title : Polusinteticheskie penitsilliny ampitsillin i oksatsillin i ikh rol' v razvitii

professional'nykh allergicheskikh dermatozov.]. *Vestnik Dermatologii i Venerologii*, 3:74–78, 1978.

[411] Price S. M. and Shupack J. L. Allergic contact dermatitis due to N,N-dimethyl-para-phenylenediamine in bacteriology technicians. *Cutis*, 21(3):330–332, 1978.

[412] I. Dahlquist. Allergic reactions to apomorphine. *Contact Dermatitis*, 3:349–350, 1977.

[413] Wozniak K. D. and Hamm G. Allergic contact eczema and vitiliginous depigmentations caused by paratertiary butylphenol. [german title : Allergisches kontaktekzem und vitiligoartige depigmentierungen durch paratertiares butylphenol.]. *Berufs-Dermatosen*, 25(6):215–219, 1977.

[414] Rothe A., Yousif S. H., and Zschunke E. Allergic contact ecezma from sodium amidotrizoate-a radiopaque substance in angiography and renography. *Contact Dermatitis*, 3(5):284–286, 1977.

[415] Malten K. E. Letterflex photoprepolymer sensitization in newspaper printers due to penta erythritol tetrakis 3 mercaptopropionate and 3 mercaptopropionic acid. *Contact Dermatitis*, 3(5):257–262, 1977.

[416] Keczkes K. and Brown P. M. Hexahydro, 1,3,5, tris (2-hydroxyethyl) triazine, a new bacteriocidal agent as a cause of allergic contact dermatitis. *Contact Dermatitis*, 2(2):92–98, 1976.

[417] Pedersen N. B. and Fregert S. Occupational allergic contact dermatitis from chloracetamide in glue. *Contact Dermatitis*, 2(2):122–123, 1976.

[418] Nethercott J. R. and Lawrence M. J. Allergic contact dermatitis due to nonylphenol ethoxylate (nonoxynol-6). *Contact Dermatitis*, 10(4):235–239, 1984.

[419] Todaro A. and Nava C. Allergic disease caused by squaric acid dibutyl ester: description of 2 cases in a work environment. [original italian title : La patologia allergica da dibutil-estere dell'acido squarico: descrizione di due casi in ambiente lavorativo.]. *Medicina del Lavoro*, 82(3):276–279, 1991.

[420] Roed-Petersen J., Batsberg W., and Larsen E. Contact dermatitis from naphthol AS. *Contact Dermatitis*, 22(3):161–163, 1990.

[421] Dannaker C. J. Allergic sensitization to a non-bisphenol A epoxy of the cycloaliphatic class. *Journal of Occupational Medicine*, 30(8):641–643, 1988.

[422] Kanzaki T. and Sakakibara N. Occupational allergic contact dermatitis from ethyl-2-bromo-p-methoxyphenylacetate. *Contact Dermatitis.*, 26(3):204–205, 1992.

[423] Lachapelle J. M. and Tennstedt D. Occupational soap dermatitis: contact allergic reaction to lauryloxypropylamine. *Contact Dermatitis*, 1(4):260, 1975.

)